AD-A046 162 HARVARD UNIV CAMBRIDGE MASS DIV OF ENGINEERING AND --ETC F/G 12/2
SCHEDULING DIAL-A-RIDE TRANSPORTATION SYSTEMS: AN ASYMPTOTIC AP--ETC(U)
SEP 77 D M STEIN N00014-75-C-0648

UNCLASSIFIED TR-670 NL

1 OF 2
AD
A046162

Office of Naval Research

Contract N00014-75-C-0648 NR-372-012

National Science Foundation Grant ENG-76-11824

SCHEDULING DIAL-A-RIDE TRANSPORTATION SYSTEMS: AN ASYMPTOTIC APPROACH

By

David M. Stein

September, 1977

Technical Report No. 670

Division of Engineering and Applied Sciences
Harvard University Cambridge, Massachusetts

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (14) TR- Technical Report No. 670	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER (9)
4. TITLE (and Subtitle) (6) SCHEDULING DIAL-A-RIDE TRANSPORTATION SYSTEMS: AN ASYMPTOTIC APPROACH.		5. TYPE OF REPORT & PERIOD COVERED Interim Rep[redacted]t.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) David M. Stein		8. CONTRACT OR GRANT NUMBER(s) (15) N00014-75-C-0648 NSF-ENG-76-11824
9. PERFORMING ORGANIZATION NAME AND ADDRESS Division of Applied Sciences Harvard University Cambridge, Massachusetts 02138		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE (11) Sep[redacted] 1977
		13. NUMBER OF PAGES 148
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) (12) 142p.		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted by the U.S. Government.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Travelling Salesmen Problem
Scheduling
Combinatorial Optimization
Dial-a-ride
Assignment Problem

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

There has been some recent interest in the "Dial-a-Ride" proposal for innovative public transportation systems. These schemes ideally provide large numbers of passengers with personalized service. A passenger is taken from his origin to his destination by a small bus; the bus may deviate en-route to collect and deliver additional passengers. In such systems the central mathematical problem is one of scheduling: the assignment of a passenger to a bus and the determination of a time for the trip.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

163 750

Our investigation is an analytic one into some of the basic aspects of dial-a-ride algorithms. We derive a class of algorithms for which performance can be measured in a certain precise sense. This sense is asymptotic (in the number of passengers who require service) and probabilistic (so that it is likely that on any given day an algorithm will perform well). The approach taken is one of preplanning at a global level. Very loosely, it is based on the following principle. Even though each passenger is unique, with his own required origin, destination and time-of-delivery, in a large system, where there are a large number of passengers, the behavior patterns of the set of all passengers can be predicted quite closely.

To analyse the problem theoretically it is necessary to abstract the essentials and to consider an idealized version. Models of the dial-a-ride problem are presented that focus upon its combinatorial nature. They study separately static version--in which demands for service are all available at the start--and dynamic versions--in which the demands arise over time. For these models, we obtain 'asymptotically optimal' algorithms that minimize simple distance or average flow-time criteria, and we evaluate suboptimal schemes.

Despite the idealization, many qualitative insights ensue. We obtain the implications of the theoretical results for the real problem and propose an approach towards the design of dial-a-ride systems: this approach is based upon the partition of the region into subregions and the specialization of buses to demand-types. The approach is of interest then, because it can be theoretically justified, and also because it has many attractive practical features. For example, resulting systems have modest computational requirements and are simple to visualize and implement. Furthermore, the techniques developed provide an analytical tool for use in the design process. It is possible to investigate changes in performance when parameters (e.g., the number of buses, the size of the region and many others) are varied. Proposed schemes can be easily evaluated without the need to resort, at a basic level, to simulation.

Office of Naval Research

Contract N00014-75-C-0648 NR-372-012

National Science Foundation Grant ENG 76-11824

SCHEDULING DIAL-A-RIDE TRANSPORTATION SYSTEMS: AN ASYMPTOTIC APPROACH

By

David M. Stein

Technical Report No. 670

September 1977

The research reported in this document was made possible through support extended the Division of Applied Sciences, Harvard University by the U. S. Army Research Office, the U. S. Air Force Office of Scientific Research and the U. S. Office of Naval Research under the Joint Services Electronics Program by Contract N00014-75-C-0648 and by the National Science Foundation under Grant ENG 76-11824.

Division of Applied Sciences

Harvard University · Cambridge, Massachusetts

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

This report is, in essence, concerned with scheduling theory. The concern takes two forms. First, there is a practical engineering scheduling problem that needs to be solved. To study this problem we have drawn upon a new theoretical approach to scheduling. And, this is our second concern, the development of a theory that indeed has practical implications.

The problem we address is that of scheduling dial-a-ride transportation systems. A solution would be the development of a methodology for analysing and designing the scheduling algorithms. Whilst aiming at a solution, we investigate to what extent a recent asymptotic probabilistic technique for the solving of hard combinatorial optimization problems is of real interest. This requires that a mathematical result be generalized for a number of problem formulations.

1.1 The Dial-a-Ride Problem

During the past decade there has been some interest in the planning of innovative public transportation systems. An area which has received substantial attention is that of 'demand responsive transportation' [19].

One of the outcomes of this research has been the 'dial-a-ride' proposal. A dial-a-ride transportation system is somewhere in the range between a rigid bus system and a flexible taxicab system, and ideally provides large numbers of passengers with personalized service. Passengers request service - to be taken from an origin to a destination - by telephone. At this time of request, an estimate of the time of collection and time of delivery is quoted. Small buses travel about the region, collecting and delivering passengers; during his trip, a particular passenger may be transferred to another bus. At any time a bus may have many passengers aboard.

The central mathematical problem in a dial-a-ride transportation system is one of scheduling: an algorithm is required that will decide to which bus (or sequence of buses) a particular passenger should be assigned and when his trip should take place. A number of algorithms have been proposed to do this scheduling. Most of these are computer aided; the computer is used for the detailed on-line decision making or as an information source, evaluating the system 'state'.

We shall be concerned with dial-a-ride scheduling algorithms at a fundamental level, avoiding many of the modifications and adjustments that must, of course, be made in an actual implementation (we shall discuss some of the practicalities in Chapter IV). There are two fundamentally very different approaches to scheduling which we shall consider. Let us describe the skeletons of these approaches.

The first is a system developed at M.I.T. by Wilson et. al. [21,22]; experimental versions have been implemented since 1972, and currently one is being tested at Rochester, New York. The underlying trait of this scheme is its search procedure for allocating passengers to buses. At each instant of time each bus has associated with it a 'prospective route', given by an ordered sequence of future stops (which may be origins or destinations(1)), and an estimated time-of-arrival at each of these stops. With each stop too there is a 'latest time-of-arrival' - the time which has been quoted to the associated passenger. A new incoming passenger must be allocated to a bus. This involves inserting his origin and destination into the prospective route of one of the buses. For each new demand, the scheduler searches through every possible insertion on every one of the buses and chooses the best, so as to

---

(1) This scheme requires no transfers for passengers: each passenger must be delivered by the same bus that collects him.

minimize a certain criterion function[2]. Basic to the design of algorithms here is the choice of the criterion function, the method of quotation, and the travel-time prediction. It appears that the procedure is very sensitive to these choices, as well as to variations in the parameters [21,22,18].

At the other extreme is a very different approach towards the scheduling. It is described by the following, which we shall refer to as the 'Michigan Scheme' since a related system has been evolving at Ann Arbor, Michigan [17]. The region in which the system operates is partitioned into a number of subregions, $r_1, r_2, \ldots r_m$. In each region r_i there is a bus that travels only in r_i. A

---

(2) The criterion is a function that concerns both present and future passengers. Present passengers are interested in their wait time, w, ride-time, r, total travel time, w+r, pickup-time deviation, Dp, and delivery-time deviation, Dd (these are deviations from the quoted times). Future passengers are acknowledged by taking into consideration the increase in tour-length, DT. Thus, for each possible assignment one can evaluate

$$a_1f_1(w) + a_2f_2(r) + a_3f_3(w+r) + a_4f_4(Dp) + a_5f_5(Dd) \qquad (1)$$

for every current passenger, and a similar value for the passenger being assigned. Here, the a_i's are weighting parameters and the f_i's are functions (generally, $f_i(x) = x$ or x^2 for all i). Summing (1) for all the passengers and adding also

$$bDT$$

where b is another weight, yields the criterion function to be minimized.

The time to a point P which is quoted a passenger is determined by

$$c.E(t_P) + d$$

where $E(t_P)$ is the expected direct travel-time to P and c and d are fixed parameters. In some versions of the algorithm the quoted times are retained as hard constraints, which must never be violated by subsequent assignments.

larger 'line-haul' bus connects the regions by travelling along a fixed route, stopping at certain 'transfer points', $p_1, p_2, \ldots p_m$ in each region. This is shown schematically in Figure 1.1.

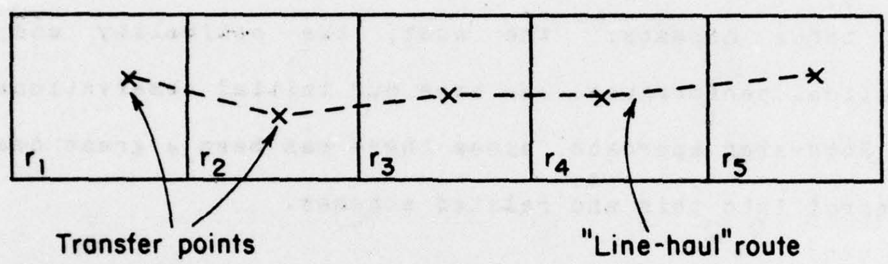


Figure 1.1 The "Michigan Scheme"

A passenger requiring to go from region r_i to region r_j is collected by the bus in r_i, transfers onto the line-haul bus at p_i, alights at p_j and is delivered by the bus in r_j. The regional buses visit their transfer points every 15 minutes (say). The regions are small enough that drivers can quite easily choose an optimal or near-optimal path between visits to their transfer point. The times which are quoted to passengers for collection and delivery are simple estimates of travel time given the load(3). Under a heavily loaded system a passenger may have to wait his turn before being collected, as the regional

(3) These quotations are not used in the scheduling procedure.

buses can make only a limited number of stops between visits to the transfer point.

We do not intend to give here a detailed evaluation of present dial-a-ride algorithms - the reader is referred to [10] - but some critical comments are in order. These focus on three aspects: the cost, the optimality and the practical performance. We base our initial observations on the Rochester approach, since there has been a great deal of research into this and related schemes.

Basic to the Rochester approach is a long and hard search. Typically, the computational effort required for such searches grows exponentially with the problem size (see the comments in Section 2.1), and we should expect it to be very expensive.

A scheme might be justified if a certain performance level can be guaranteed. The Rochester search is a local one [5], both with respect to time and with respect to the passengers. That is, only a single passenger is assigned (or at best only a small set of passengers is reassigned) when the schedule is updated. Further, optimality (with respect to the criterion function) is ensured for only that point in time. For local searches it is very rare that good performance can be guaranteed. Recent research in [22] has focused on the criterion function. However, there is no guarantee that the performance of the system as a whole (measured, say, in terms of average travel-time and/or

variance of travel-time(4)) will be improved if the utility functions of all current passengers are maximized whenever decisions are made. This is true even if reasonably acceptable utility functions for the passengers could be represented, itself a notoriously messy problem.

Lastly, what about its practical performance? In the early stages of implementation, particularly when the system was heavily loaded, there were often roundabout devious routes for passengers, with resulting customer dissatisfaction [18]. In general, it appears that the results on performance are inconclusive.

In contrast, the simplicity of the scheme at Michigan is very attractive. The scheduling cost is relatively low, since the computer is used only to store the demands and access them efficiently. (Optimal tours for the regional buses might also be determined on-line by a computer. Then each tour is but a small, simple 'travelling salesman tour' - see Chapter II - where the distances between points satisfy the triangle inequality. For these problems there are some recent efficient heuristic algorithms [12].)

In practice the Michigan scheme has been found to work well [17]. But there are still some vague questions. Can one justify the use of the Michigan approach with respect to some optimality criterion; and how should one design the regions, the fixed routes and the timing of transfers?

---

(4) These were the measures loosely used in [22].

1.2 Our Approach

Our investigation will be an analytic one into some of the basic aspects of dial-a-ride algorithms. We shall derive a class of algorithms - or, more correctly, an approach towards the design of algorithms - for which optimality can be measured in a certain precise sense. This sense is asymptotic (in the number of passengers who require service) and probabilistic (so that the probability is high that on any given day the algorithm will perform well).

The approach taken is one of preplanning at a global level. Very loosely, it is based on the following principle. Even though each passenger is unique, with his own required origin, destination and time-of-delivery, in a large system, where there are a large number of passengers, we can predict quite closely the behaviour patterns of the set of all passengers. This is the 'equalizing' effect of the law of large numbers that has been observed in many physical phenomena, the classic example being the thermodynamic principles for the behaviour of gases.

To analyse the problem theoretically it is necessary to abstract the essentials and to consider an idealized version. Chapters II and III present an analysis of models of the dial-a-ride problem that focus upon its combinatorial nature. They study separately static versions - in which demands for service are all available at the start - and dynamic versions - in which the demands arise over time.

For the models, we obtain 'asymptotically optimal' algorithms that minimize simple distance or average flow-time criteria, and we evaluate suboptimal schemes.

Despite the idealization, many qualitative insights result, and in Chapter IV, we return to the real problem. Here, drawing upon the theoretical results, we obtain their implications for the real problem and propose an approach towards the design of dial-a-ride systems. The approach is of interest then, because it can be theoretically justified, and also because it has many attractive practical features. Furthermore, the techniques developed in Chapters II and III provide us with a powerful analytical tool for use in the design process. We can investigate changes in performance when parameters (e.g. the number of buses, the size of the region and many others) are varied. Proposed schemes can be easily evaluated without the need to resort, at this basic level, to simulation.

It is interesting to note, at this point, that the 'Michigan Scheme' described above can be considered as belonging to our class of algorithms. Thus, we are able to say in what sense it is optimal as well as to gain some valuable insights towards its improvement. The attractive simplicity of the Michigan scheme is common to all of our proposed algorithms.

CHAPTER II

STATIC BUS PROBLEMS

In this chapter we begin a theoretical analysis of dial-a-ride bus systems. In Section 2.0, we introduce the idealized 'bus problems' which we shall be developing, and state the basic assumptions and terminology. Our approach is based on an interesting theorem of Beardwood, Halton and Hammersley [2] and on a recent paper of Karp [9]; we describe these results within the context of combinatorial optimization problems in Section 2.1. Sections 2.2 and 2.3 present a theory for single- and multiple-bus problems.

Throughout this chapter the presentation is informal, with only heuristic justifications and proofs. Detailed proofs of the results are delegated to an appendix.

2.0 Idealized Bus Problems

The travelling salesman problem (TSP) - see the definition below - has received much attention in the operations-research literature. Not least among the reasons for this is the fact that the TSP can be regarded as a prototype for many realistic problems. Clearly, there is a nontrivial relationship between the TSP and the dial-a-ride

problem, and in this chapter we shall exploit this relationship. First let us describe the problems to be considered, and state the basic assumptions and our terminology.

We are given a bounded region R in the plane, with area a. Demands arise in R. A demand p = (o,d) is a pair of points in R with origin o and destination d. A set of m buses is at our disposal for meeting these demands - i.e. for visiting the set of points. A tour for a bus is the sequence of points which it visits. Before a bus can visit a destination the relevant passenger must be on board, and so the corresponding origin must already have been visited. We refer to this as the 'feasibility constraint'; tours which satisfy it are termed 'feasible'.

We distinguish between static and dynamic versions of the problem. In the static version we are given, at time t=0, a collection of n demands and we are to devise feasible tours for the buses so that all 2n points are visited. In dynamic problems, to be discussed in Chapter III, the demands arise as time progresses according to some random process. Again feasible tours are to be devised (of course a point can be visited only after the demand arises). The problem faced is to devise tours so that a certain criterion is minimized. The criterion will usually be a function of time or of distance travelled.

We assume that the buses are all of infinite size. The distance between two points in R is the euclidean length of the straight line joining them, so if R is not convex the tour might leave R. The buses travel at unit speed, and no time is wasted when a passenger embarks or alights. Also, transfers are instantaneous (see Section 2.3). These assumptions will be discussed further in Chapter IV.

An instance of a static bus problem of size n is specified by a set of n demand pairs. Our approach is probabilistic and we must define a probability distribution from which the problem instances are drawn. We assume (for simplicity, although this can be considerably weakened) that all origins and destinations are drawn independently from the uniform probability distribution over R. We refer loosely to this fact by saying that a particular problem instance (of size n) is 'random'.

We shall also be making use of the well-known travelling salesman problem. An instance of this problem (the euclidean version) is given by a set of n points within R. We are required to find a path[1] which passes through these n points and which has shortest length. In a 'random' instance of the TSP, the n points are drawn independently from a uniform probability distribution over R.

---

(1) By a 'travelling salesman problem' we mean this 'open' version, in which the required tour through the n points need not be closed.

In the algorithms to be considered, the region R is partitioned into subregions (this emphasis will be justified). When we say a bus 'visits' certain points in a subregion we mean that it enters the subregion and performs an optimal travelling salesman tour on the designated points there. When the bus then visits points in another region, it travels to the closest of the new points and similarly performs an optimal travelling salesman tour in that region.

If we think of R as being divided into m equal subregions each of area a/m, then any given demand pair originates in any particular region with probability $1/m$, and has destination in any region with probability $1/m$. If there are n demands, with n a large number, then with high probability there are approximately n/m origins and n/m destinations in each region. This follows from the strong law of large numbers [14]. However, we might loosely assume that there are exactly n/m origins and n/m destinations in each region. This is an example of the sort of imprecision that exists in the following pages. Rigorous statements and detailed proofs of the results are collected in the appendix, so we are free to concentrate here on the spirit of the ideas. It should be kept in mind throughout that inexactitudes as that above hold only asymptotically (in n) with probability 1.

Our analysis will begin with a very simple bus problem, and the results will then be extended to more complex problems with corresponding increased realism. First, however, we digress to describe the aspects of scheduling theory upon which our approach is based.

2.1 Combinatorial Optimization and the Travelling Salesman Problem

Recently there have been some theoretical advances in the understanding of combinatorial optimization problems. Most important has been the acknowledgement that there exists a class of 'hard problems' (NP-complete problems). These are hard in the sense that the computing time required to find their solution by any known algorithm explodes exponentially as the size of the problem increases. For these hard problems, then, there seems to be no way to avoid, essentially, the enumeration of a very large number of possible alternatives. This is inefficient; an 'efficient' algorithm would be one with an execution time behaving as p(n), a polynomial function of n. These notions have been made mathematically precise - see Aho et. al. [1], Coffman [5].

For example, the TSP has been shown to be NP-complete [16], and there are no known efficient algorithms for its solution. To solve it, we need to evaluate very many of the

n/2! possible orderings of the points, or tours. Furthermore, the existence of an efficient algorithm for the TSP would imply, and be implied by, the existence of efficient algorithms for a whole class of hard scheduling and other combinatorial optimization problems. At present, this appears extremely unlikely.

This is indeed bad news, but practical problems must still be solved. To this end, note that we required above that the algorithm guarantee the optimal solution. An easier question which might be asked would be one with the optimality requirement relaxed. So, find an efficient algorithm A (if one exists!) that will provide a possibly nonoptimal cost - call it c(A) - but one that is close to the optimal cost, c*. More precisely, an algorithm A is said to solve the problem to within the ratio y (y is a real number, larger than 1) if

$$c(A) \leq yc^* \tag{2.1}$$

(We shall also use the terminology: A is ε-optimal if

$$c(A) \leq (1 + \varepsilon)c^* \quad)$$

This approach has been found to be very useful for certain special problems, and of no use at all for certain others. For the travelling salesman problem the best that has been achieved in this sense is an algorithm (Christofides [3]) that solves the problem in polynomial time to within 3/2.

This ratio is still a little high and one might expect to do better. For, given an instance of the travelling salesman problem, it is not too hard to draw a path which looks pretty much the shortest, and one might then be tempted to say that with high probability it is close to the optimum. This implies that we should relax not only the optimality requirement but also the guarantee which the algorithm promises. This notion coincides with a very recent approach, suggested by Karp [9]. Suppose that the problem instance is derived from a certain known probability distribution. Can we then produce an efficient algorithm that will perform well (i.e. to within an acceptable ratio) with high probability? *This probabilistic* approach can best be illustrated via the TSP.

Let us suppose that a particular problem instance of the TSP is chosen by drawing n points independently from a uniform probability distribution over the planar region R. Let L_n be the length of the shortest path through these n points. Then L_n is a random variable.

The following theorem(2) is due to Beardwood et. al. [2].

---

(2) The problem studied in [2] is actually the <u>closed</u> TSP in which a closed tour is sought. It is easy to see that the theorem still holds for the open TSP.

Theorem 1

If the region R is bounded and has area a, then there exists an absolute constant b such that

$$\lim_{n\to\infty} \frac{L_n}{\sqrt{n}} = b\sqrt{a} \quad \text{almost everywhere (a.e.)} \qquad (2.2)$$

[]

The existence of the limit in (2.2) is part of the assertion. The constant b has been estimated by Monte Carlo experiments to be .75. Of course, the rate of convergence of the sequence $L_n/\sqrt{n}$ is important; this aspect of the problem will be discussed in detail in Section 4.3. We shall also discuss there the case in which the probability distribution is not uniform.

By Theorem 1, then, the length of the shortest path, a random variable, is asymptotically (i.e. for n large) equal to $b\sqrt{a}\sqrt{n}$ with probability 1. The value $b\sqrt{a}\sqrt{n}$ is a non-random function of n and so for large n there is no distinction between the random and non-random versions of the problem, and we can predict with probability 1 the length of the optimal tour through any random points. In particular, suppose that an algorithm A yields a tour length of L_n^A through n random points, and that n is very large. Then, by definition (2.1), this algoritm solves the problem to within $L_n^A/b\sqrt{a}\sqrt{n}$ with probability 1. It is in this asymptotic probabilistic sense that we are able to investigate the optimality of various algorithms.

Consider the following algorithm (similar to that of Karp [9]) which yields a path for any instance of the TSP.

Algorithm 1

Divide the region R into m subregions each of area a/m, and label them $r_1, r_2, \ldots r_m$. Using an optimal algorithm, construct an optimal travelling salesman tour within each of these regions individually. Now, visit the regions $r_1, r_2, \ldots r_m$ in order; upon completing a tour in region r_i, visit the closest unvisited point in region r_{i+1} (take $r_{m+1} = r_1$) and thence traverse the tour in that region. []

For n large enough there will be n/m points in each region; the tour within each region has length(3) (by Theorem 1)

$$b\sqrt{\frac{n}{m}}\sqrt{\frac{a}{m}} = \frac{b}{m}\sqrt{n}\sqrt{a} \qquad \text{(a.e.)}.$$

Calling the total distance T_n^K we have

$$T_n^K \leq m\frac{b}{m}\sqrt{n}\sqrt{a} + m\Delta \qquad \text{(a.e.)}$$

---

(3) It should be clear that these equations and inequalities hold in the asymptotic probabilistic sense only. The derivation can easily be made rigorous. The appendix of Chapter II will illustrate how this can be done.

where Δ is the diameter[4] of the region R.

Hence,

$$\frac{T_n^K}{\sqrt{n}} \leq b\sqrt{a} + \frac{m\Delta}{\sqrt{n}} \qquad \text{(a.e.)}$$

and, given any ε, for n large enough,

$$\frac{T_n^K}{\sqrt{n}} \leq b\sqrt{a} + \varepsilon \qquad \text{(a.e.)}$$

So, loosely speaking, Algorithm 1 is 'asymptotically optimal'. More formally, we have

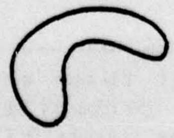
<u>Corollary 1</u>

Given any $\varepsilon>0$, there exists an $N(\varepsilon)$ such that for any random problem of size n, with $n \geq N(\varepsilon)$, Algorithm 1 is ε-optimal with probability 1. []

It can further be easily shown - see Karp [9] - that if m grows as $\log\log n$ (so that $m/n \to 0$ as $n \to \infty$) and an $O(n.2^n)$ algorithm (e.g. that in [8]) is used within each subregion, then Algorithm 1 runs in time $O(n\log n)$ a.e.

---

(4) The diameter of R is $\min\{\|x-y\|;\ x,y \in R\}$. The term involving the diameter becomes negligible in the limit. The formulas are independent of the shape (or convexity) of the region. For example, in a region with shape



the travelling salesman tour will still be hardly likely to exit from the region if n is large enough.

2.2 Static, Single-Bus Problems

2.2.1 Optimal Algorithms

The simplest static bus problem is one in which there is only a single bus. We wish to choose a feasible tour through the n demand pairs so as to minimize the total distance travelled (this will be the same as minimizing the time by which the final passenger is delivered). This problem differs from the TSP only in the feasibility constraint.

As a first observation, note that the problem is NP-complete (see Observation 2.0 in Appendix II).

Let us define(5)

L_n = length of the optimal travelling salesman tour through n random points in R.

Y_n = length of the optimal feasible bus tour through n random demand pairs in R.

Observe that, by dropping the feasibility constraint, we would obtain a tour of length L_{2n}. Hence

$$L_{2n} \leq Y_n \qquad (2.3)$$

An upper bound to Y_n can be obtained by any suboptimal algorithm which might be suggested. For example, if we first visit all origins in R and then visit all destinations, we would obtain a tour with length $L_n^1+L_n^2$,

---

(5) Both L_n and Y_n are random variables.

where L_n^1 and L_n^2 are the lengths of the two travelling salesman tours on n random points each. This yields

$$Y_n \leq 2L_n.$$

We can easily improve on this upper bound as follows. Divide R into two equal subregions of area $a/2$ each, calling them regions r_1 and r_2. First, visit all origins in r_1 (there will be $n/2$ points, from demands of the form (r_1,r_1) and (r_1,r_2) for n large). Second, visit the origins and the destinations from demands (r_1,r_2) in r_2 (there will be $3n/4$ points here). Third, visit r_1 again, visiting the remaining points there - these will be the $n/2$ destinations from demands (r_1,r_1) and (r_2,r_1). Finally, visit the last destinations in r_2 - there will be $n/4$ points from demands (r_2,r_2). For n large this tour has approximate length (a.e.)

$$b(\sqrt{n/2} + \sqrt{3n/4} + \sqrt{n/2} + \sqrt{n/4})\sqrt{a/2}$$
$$\simeq 1.96b\sqrt{a}\sqrt{n} = 1.96L_n.$$

The obvious generalization is easy, and we can give the following algorithm, producing a feasible tour which we call T_m^o.

<u>Algorithm 2</u> (producing the tour T_m^o)

Partition the region R into m equal subregions, each of area a/m, and label them $r_1,r_2,\ldots r_m$. Visit the regions in numerical order; in each region r_i collect all the origins as well as the destinations from regions $r_1,r_2,\ldots r_{i-1}$. This is the 'first passage' through the

regions. Following this, again visit the regions $r_1, r_2, \ldots r_m$ in order, visiting all the remaining destinations in each. This is the 'second passage' through the regions. Each time a region is visited, the tour is a travelling salesman tour. []

It is shown in Lemma 2.2 that the tour T_m^o has length

$$\frac{4}{3}\sqrt{2}b\sqrt{a}\sqrt{n} + \sqrt{n}O(1/m) \qquad (2.4)$$

If we now assume, as appears reasonable at this stage, that there exists an absolute constant c such that

$$\lim_{n\to\infty} \frac{Y_n}{\sqrt{n}} = c\sqrt{a} \qquad \text{(a.e.)} \qquad (2.5)$$

then (2.4) and (2.5) (noting that $L_{2n} = \sqrt{2n}b\sqrt{a} = \sqrt{2}L_n$) together give

$$\sqrt{2} \leq c \leq 4\sqrt{2}b/3.$$

More difficult is Lemma 2.3 where it is shown that the value of $4\sqrt{2}b/3$ is also a lower bound to c (the rigorous proofs in the appendix do also guarantee the existence of the limit in (2.5) and hence too that c exists). We give a rough outline to the proof of Lemma 2.3. Given any optimal tour with length Y_n, it is possible to construct a suitable division of the area R into m equal subregions (for any $m\geq 1$) such that the following infeasible tour does yield a lower bound to the value Y_n. This infeasible tour visits the regions successively, twice each, in exactly the manner of the tour T_m^o, except that the destinations of the

form (r_i, r_i), $i=1,2,\ldots m$, are visited in each region r_i on the first passage through the regions and not on the second[6]. Letting $m \to \infty$ the lengths of the infeasible tours converge to $4\sqrt{2}b\sqrt{n}/3$, a.e.

From Lemmas 2.2 and 2.3 we can state

<u>Theorem 2</u>

$$\lim_{n\to\infty} \frac{Y_n}{\sqrt{n}} = \frac{4}{3}\sqrt{2}b\sqrt{a} \qquad \text{a.e.}$$

$$= 1.89b\sqrt{a}. \qquad []$$

Henceforth for convenience we write

$$c = 4\sqrt{2}b/3.$$

We shall also refer to the 'optimal tour T^o' with length $c\sqrt{n}\sqrt{a}$, where $T^o = \lim_{m\to\infty} T^o_m$; by this rough statement we mean that we can approach arbitrarily close to the optimal value of Y_n by a tour T^o_m given by Algorithm 2; this is still asymptotic in n (with $n/m \to \infty$) a.e.

More precisely, we get from Lemma 1,

---

(6) Note that if there are n_i points to be visited in a certain region r_i and if n_i' and n_i'' are visited on the first and second passages through r_i respectively, then the tour in r_i has length behaving as

$$y_i = b\sqrt{a/m}(\sqrt{n_i'} + \sqrt{n_i''})$$

with $n_i' + n_i'' = n_i$

The value y_i is minimized if $|n_i' - n_i''|$ is made as large as possible.

<u>Corollary 2</u>[7]

Given any $\varepsilon>0$, there exist $N(\varepsilon)$ and $m(\varepsilon)$ such that for $n\geq N$, the tour T_m^o given by Algorithm 2 minimizes the total distance travelled to within $1+\varepsilon$, with probability 1. []

<u>2.2.2 Comments on the Tour T_m^o</u>

<u>(i) Uniqueness of the Tour T^o</u>

Note that the tour T^o described above does not uniquely solve the minimum distance problem. Consider a tour $T_m^{o'}$ which visits the regions $(r_1, r_2, \ldots r_m)$ in the same manner as T_m^o on the first passage through the regions. On the second passage, however, the regions are visited in the order $(r_m, r_{m-1}, \ldots r_1)$, and once again all feasible destinations are visited in each region. Clearly T_m^o and $T_m^{o'}$ have the same length (asymptotically a.e.) and

$$T^{o'} \stackrel{\Delta}{=} \lim_{m\to\infty} T_m^{o'}$$

would also solve the problem.

---

(7) This statement can be strengthened slightly: Given any $\varepsilon>0$ there exists an $M(\varepsilon)$ and a function $N(m,\varepsilon)$ such that for all $m\geq M$, the tour T_m^o minimizes the distance to within $(1+\varepsilon)$ a.e., whenever $n\geq N(m,\varepsilon)$.

In later sections, particularly for the dynamic problem, we shall be interested in other criteria; there is then an inefficiency in the tour $T_m^{o'}$ as it is 'tail-heavy' with passengers having on the average a longer travel-time. This justifies our present emphasis on the tour T_m^o.

There is yet another tour which minimizes the distance travelled. Consider visiting regions $(r_1, \ldots r_m)$ in the same order on two separate passages. On the first, collect in each region r_i all origins of the form (r_i, r_j), $j \leq i$. On the second passage, collect all origins of the form (r_i, r_j), $j > i$ and simultaneously deliver all feasible destinations (these will be all of the destinations in r_i). It is easy to see that this tour will have the same length as the tour T_m^o, but will be even more tail-heavy than was $T_m^{o'}$.

We might emphasise the easy principle underlying the optimal tour T_m^o, which will later be seen to be important. Visit the regions successively; each time the bus exits from a region choose as next region the one with the most unvisited feasible points (break ties arbitrarily), and visit all of these points.

__(ii) Simple Tours__

This is a suitable juncture at which to digress to make a further important observation on the tour T^o. It belongs to a class of tours which we shall term 'simple'.

Conceptually, a simple tour is one that visits the points in R region-by-region. The tour segment in r_i can be constructed at the time at which the bus enters r_i, in that all points visited in r_i were feasible at this time of entry.

More precisely, let m be any integer, $m \geq 1$. Define S_m, the set of 'simple tours on m subregions' as follows. A feasible tour T belongs to S_m if there exists a partition of R into at least m subregions of area a/m each, with the two properties:

1) if T enters a region r_i at time t_i, then the points which are visited in r_i were already feasible at time t_i,

2) if T visits p points in r_i, then we can assume that these p points are randomly distributed over r_i (and that T performs an optimal travelling salesman path on these p points).

Now define

$$S = \bigcup_{i=1}^{\infty} S_i$$

Then any tour in S is feasible, and

$$S_i \subset S_{i+1} \subset S \qquad \text{for any } i=1,2,\ldots$$

We call S the set of 'simple tours'.

By Corollary 2, for the single-bus static problem it is enough to consider only simple tours. Most tours which we encounter will be simple; in Chapter III we shall restrict our consideration to simple tours only.

2.2.3 Interesting Suboptimal Algorithms

It is of interest to describe briefly some suboptimal algorithms. These are closely related to currently used techniques and will be needed subsequently.

The following might be considered to be the 'Michigan Algorithm', described in Chapter I, for the static single-bus case.

Partition the region R into m subregions. Calculate an optimal travelling salesman tour on the origins within each region. Visit all regions, performing this optimal tour in each, and linking the regions together using some fixed-route, of any length g. Then, again return along the same fixed-route, deviating within each region to deliver all destinations, using an optimal travelling salesman tour within each region.

Let Y_n^M be the total length of this bus tour.

Then

$$Y_n^M = 2mL_{n/m} + 2g$$

and

$$\lim_{n\to\infty} \frac{Y_n^M}{\sqrt{n}} = 2b\sqrt{a} \quad \text{a.e.}$$

The distance g travelled by the fixed-route bus is asymptotically negligible. It will henceforth be convenient to think of the region R as the interior of a circle, and the subregions as sectors (see Figure 2.1).

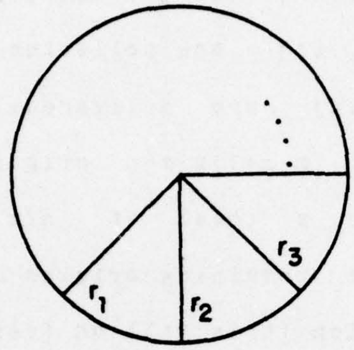


Figure 2.1 Partition of R into subregions

It is then easy to visualize, for example, a bus visiting the subregions in any required order, and to eliminate the need for a fixed-route bus. However, this convenience is not necessary.

The 'Michigan tour' described above is now seen to be only a very simple extension of Algorithm 1 to the bus problem. With respect to distance travelled, it can be considered the same as the tour T_1^o given by Algorithm 2.

Most important, from the practical point of view, is that (see (2.1)) the algorithm solves the problem to within $2b/c = 1.06$ a.e.

A second suboptimal algorithm is the following 'fixed-route' algorithm. Again the region is divided into m subregions and a priori it is specified that the regions will be visited first in the order $r_1, \ldots r_m$ and then in

the order $r_m, r_{m-1}, \ldots r_1$. On the first passage, origins of the form (r_i, r_j), $i \geq j$ are collected and destinations of the form (r_i, r_j), $i<j$ are delivered. In region r_i, there will be $(m-i+1)n/m^2$ origins and $(i-1)n/m^2$ destinations visited, a total of n/m points. On the second passage all remaining origins and destinations are visited in each region (this will be feasible).

Let Y_n^F be the total length of this bus tour. Then,

$$Y_n^F = 2mb\sqrt{\frac{n}{m}}\sqrt{\frac{a}{m}}$$

or, more precisely,

$$\lim_{n\to\infty} \frac{Y_n^F}{\sqrt{n}} = 2b\sqrt{a} \quad \text{a.e.}$$

and this algorithm also solves the problem to within 1.06 a.e.

2.3 Static, Multiple-Bus Problems

2.3.1 Introduction

We now consider the static problem in which there are k buses. The problem faced is essentially one of multiprocessing. Any passenger can be served by any bus and we are to allocate the passengers to buses in some optimal fashion.

Two possible generalizations of the previous criterion immediately arise, and we get the problems:

(i) *Minimize the total distance travelled by all* k buses (i.e. minimize fuel used), and

(ii) Minimize the time-to-completion (i.e. time-to-delivery of the final passenger).

A feasible solution to the problem is given by a set of k tours, one for each bus. For any such k tours, let

$$x_n^i = \text{distance travelled by bus } i,\ i=1,\ldots k$$

$$z_n^k = \sum_{i=1}^{k} x_n^i$$

$$y_n^k = \max\{x_n^i,\ i=1,\ldots k\}.$$

In (i) we wish to choose the tour to minimize z_n^k; let Z_n^k be the optimal such value. In (ii) we wish to minimize y_n^k; let Y_n^k be this optimal value. (Z_n^k and Y_n^k are random variables.)

Note that

$$Z_n^1 = Y_n^1 = Y_n$$

(Y_n is as defined in the last section).

It is easy to see that for any instance of the problem we can always achieve

$$Z_n^k \leq Y_n$$

- simply use only one bus and keep $k-1$ idle. It appears reasonable to claim that $Z_n^k = Y_n$. This claim is in fact verified by Lemma 3.1.

<u>**Lemma 3.1**</u>[8]

$$\lim_{n\to\infty} \frac{Z_n^k}{\sqrt{n}} = c\sqrt{a} \qquad \text{a.e.} \qquad [\,]$$

Thus, by increasing the number of buses we do not improve upon the total distance travelled or upon the fuel consumption. If instead, the final time-to-delivery of all passengers is of interest, we must investigate Y_n^k. Then we require that each of the buses absorbs part of the load, and the jobs of collecting and delivering passengers must be executed in parallel. For efficiency, all passengers should travel an equal distance, so that all buses are busy all the time. Clearly too, we must have

$$Y_n^k \geq \frac{Y_n^1}{k}$$

We ask, under what circumstances do we in fact get

$$Y_n^k = \frac{Y_n^1}{k} \quad ?$$

Theorem 3 to follow indicates that this lower bound is attainable (in the limit) if passengers may transfer[9] between buses, as in the scheme at Michigan. If passengers cannot transfer, so that the bus collecting a passenger must also deliver him, as in the scheme at Rochester, then each

---

(8) The proof of Lemma 3.1 which is given in the appendix is somewhat indirect, and an easier proof has been found to be evasive. However, the proof does identify additional problems as being of interest: these are problems in which at most i transfers are allowed during the tours of the buses, for i=1,2,3,...

(9) At a 'transfer point', at least two buses meet and the passengers on board are able to move arbitrarily between them.

bus must travel a greater distance.

2.3.2 Transfers Allowed

Theorem 3

If transfers are allowed

$$\lim_{n\to\infty} \frac{Y_n^k}{\sqrt{n}} = \frac{c\sqrt{a}}{k} \quad \text{a.e.} \qquad [\,]$$

Consider the following algorithm, producing a tour which we call $T_m^{k^o}$.

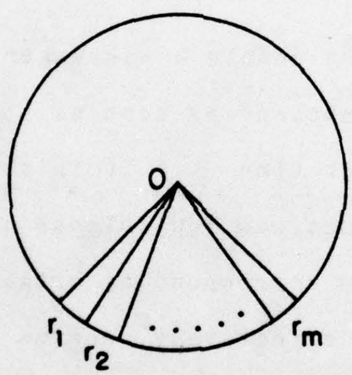


Figure 2.2 Algorithm 3

Algorithm 3 (producing $T_m^{k^o}$)

Let m be any positive integer. Partition the region R into mk subregions each of area a/km. Bus i will be visiting only regions $r_{(i-1)m+1}, \ldots r_{im}$, $i=1,\ldots k$. First, let each bus i visit region $r_{(i-1)m+1}$, collecting all origins there. Then, let all buses meet at a

prearranged transfer point (the point O in Figure 2.2) and passengers transfer onto the bus that will ultimately be visiting the region containing the required destination. Each bus i then visits region $r_{(i-1)m+2}$, collecting all origins there and delivering all feasible passengers (from $r_{(j-1)m+1}$, $j=1,\dots k$). The k buses meet at the next transfer point when passengers who have just been collected again transfer to their required buses. In this way, after m transfers, the buses will have collected all origins. Then (in the manner of Algorithm 2) bus i revisits regions $r_{(i-1)m+1},\dots r_{im}$, delivering the remaining passengers. []

The m transfers enable a passenger to be on the bus visiting his destination as soon as possible. Whenever a bus enters a region at time t, it is able to visit every destination there that was feasible at t, irrespective of which bus visited the corresponding origin. Each bus visits $2n/k$ points and effectively serves n/k passengers. Furthermore, each bus travels through $(1/k)$th of the total area a (also, each passenger needs to travel at most through an area a/k) and travels a total distance $c\sqrt{n/k}\sqrt{a/k} = c\sqrt{n}\sqrt{a}/k$. The following corollary is then obvious (but see the detailed proof of Theorem 3).

<u>Corollary 3</u>

Given any $\varepsilon>0$ there exist $N(\varepsilon)$ and $m(\varepsilon)$ such that for $n\geq N$ the tour $T_m^{k^o}$ given by Algorithm 3 is ε-optimal for problem (ii) (i.e. minimizing completion-time) with

probability 1. []

For the tours T_m^o given by Algorithm 2, we defined a 'tour' $T^o = \lim_{m\to\infty} T_m^o$. There was no conceptual difficulty in doing this: the larger m became the sooner the destination whose origin had just been collected became feasible. The tour $T_m^{k^o}$ of Algorithm 3 now requires m transfer points; in the analogous limit as $m\to\infty$, a limiting optimal tour would require continuous transfers on the first passage through the region R. We thus refrain from defining a corresponding limiting 'tour'.

If we define the 'Michigan tour with k buses' to be the tour yielding $T_1^{k^o}$ (so this tour has only one transfer) what do we lose? As before, the distance travelled by each bus before the transfer (i.e. n/k origins in an area a/k are collected) is $b\sqrt{n/k}\sqrt{a/k}$. After the transfer, n/k destinations are visited in the same area. Hence, if $Y_n^{k^M}$ is the completion time which this algorithm yields, then

$$\lim_{n\to\infty} \frac{Y_n^{k^M}}{\sqrt{n}} = \frac{2b\sqrt{a}}{k} \qquad \text{a.e.}$$

Theorem 3 gives that this tour solves problem (ii) to within $(2b/k)/(c/k) = 1.06$ a.e. once again; we lose but 6% by restricting ourselves to only one transfer.

It is possible to improve on this algorithm slightly. Passengers with both origin and destination in the same area serviced by bus i, say, might in fact be able to be delivered before the transfer point. Thus, divide the region into mk subregions, with each bus servicing m regions, and a single transfer taking place after the first passage through these regions. Now m can be made large with no conceptual difficulty[10]. Such schemes will be discussed in greater detail for the dynamic problem of Section 3.3.

2.3.3 No Transfers Allowed

We now turn our attention to schemes in which transfers are not allowed. Of course, Lemma 3.1 still holds, but Theorem 3 does not. It is easy to see that

$$Y_n^k \geq \frac{Y_n^1}{k} \tag{2.6}$$

and, for k>1, it appears that a strict inequality holds if there are to be no transfers. We do not have results as strong as Theorem 3 for this case.

---

(10) This scheme yields a completion-time, $Y_n^{k^{M'}}$ satisfying

$$\lim_{n\to\infty} \frac{Y_n^{k^{M'}}}{\sqrt{n}} = \frac{2}{3} b\sqrt{a}[(1 + 1/k)^{3/2} - (1 - 1/k)^{3/2}] \quad \text{a.e.}$$

The proof of this is similar to that of Theorem 6 (Appendix III).

(i) Two-Bus Case

For $k=2$ consider the following algorithm, which is a 'fixed-route scheme' for two buses (recall the fixed-route algorithm with one bus). Once more divide R into m subregions $r_1, r_2, \ldots r_m$. Let bus I visit the regions in order $(r_1, r_2, \ldots r_m)$ collecting and delivering all passengers with demands (r_i, r_j), $j>i$. Let bus II visit the regions in order $(r_m, r_{m-1}, \ldots r_1)$, collecting and delivering passengers with demands (r_i, r_j), $j<i$. Passengers with demands (r_i, r_i) are divided evenly between the buses and, as we can let $m \to \infty$ these passengers will not cause difficulties. (See *Figure* 2.3.)

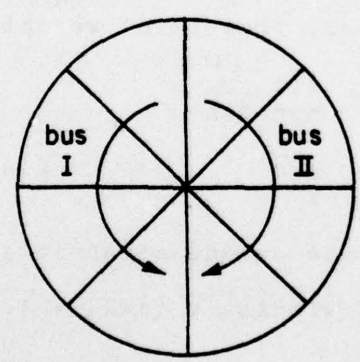


Figure 2.3 Two-bus static scheme without transfers

Each bus will visit n/m points in each region as before: for example, bus I in region r_1 will collect $(m-1)n/m^2$ origins (r_1, r_j), $j>1$, $n/2m^2$ origins

(r_1,r_1) and will deliver $n/2m^2$ destinations (r_1,r_1).

Thus, if Y_n^{2F} is the completion-time which this algorithm yields, then

$$\lim_{n\to\infty} \frac{Y_n^{2F}}{\sqrt{n}} = mb\sqrt{\frac{1}{m}}\sqrt{\frac{a}{m}}$$

$$= b\sqrt{a} \qquad \text{a.e.}$$

Having considered numerous other two-bus schemes, we are lead to conjecture that

$$\lim_{n\to\infty} \frac{Y_n^2}{\sqrt{n}} = b\sqrt{a} \qquad \text{a.e.}$$

i.e. that this scheme is asymptotically optimal (a.e.) for the minimal completion-time problem (problem (ii)) with no transfers.

Be this as it may, from (2.6) we obtain

$$\lim_{n\to\infty} \frac{Y_n^2}{\sqrt{n}} \geq \lim_{n\to\infty} \frac{Y_n^1}{2\sqrt{n}}$$

$$= \frac{c}{2}\sqrt{a} \qquad \text{a.e.}$$

so that the fixed-route scheme at worst solves problem (ii) with no transfers to within $b/(c/2) = 1.06$ a.e.

(ii) Three-Bus Case

A similar scheme which we conjecture to be optimal for $k=3$ buses is the following. (See Figure 2.4.)

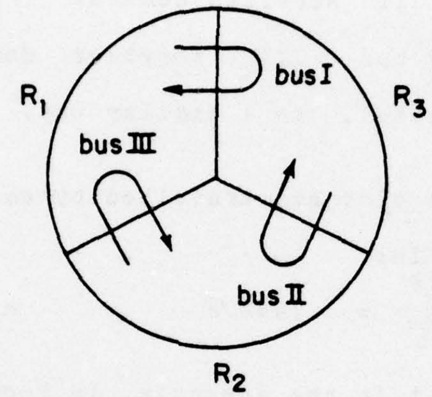


Figure 2.4 Three-bus static scheme without transfers

Divide R *into* $3m$ *subregions* $r_1, r_2, \ldots r_{3m}$. Let $R_1 = \bigcup_{i=1}^{m} r_i$, $R_2 = \bigcup_{i=m+1}^{2m} r_i$, $R_3 = \bigcup_{i=2m+1}^{3m} r_i$. Then bus I services demands (R_1,R_1), (R_1,R_2) and (R_2,R_1) by travelling in order through $(r_1,r_2,\ldots r_m)$, through $(r_{2m},r_{2m-1},\ldots r_{m+1})$ and through $(r_1,r_2,\ldots r_m)$ again. (Call these the first, second and third passages respectively.) On the first passage the bus collects in r_i, $1\leq i\leq m$, all origins of the form (r_i,r_j), $i\leq j\leq 2m$, and delivers all feasible destinations, i.e. of the form (r_j,r_i), $1\leq j<i$. On the second passage in regions r_i, $m+1\leq i\leq 2m$ (in reverse order!) origins of the form (r_i,r_j), $1\leq j<i$ and all feasible destinations (r_j,r_i), $1\leq j\leq m$ and $i<j\leq 2m$ are visited. Finally, on the third passage, all remaining destinations in r_i, $1\leq i\leq m$ are visited, i.e. (r_j,r_i), $m+1\leq j\leq 2m$ and (r_i,r_i). At the

same time, bus II services demands (R_2,R_2), (R_2,R_3) and (R_3,R_2), and bus III services demands (R_3,R_3), (R_3,R_1) and (R_1,R_3), in a similar way.

As $m\to\infty$, the distance travelled by each bus, defined to be Y_n^{3F} satisfies

$$\lim_{n\to\infty} \frac{Y_n^{3F}}{\sqrt{n}} = .798b\sqrt{a} \qquad \text{a.e.}$$

We derive this fact in the appendix, in Proposition 3.2.

(iii) k-Bus Case

The approach taken above can be used to obtain what appear to be good schemes for k buses, with k large.

Suppose there are $k=s^2$ buses. Divide R into s subregions and let each bus service a particular demand-type, i.e. (r_i,r_j), $i,j=1,2,\ldots s$. For $i\neq j$, the bus serving (r_i,r_j) will collect n/s^2 points in each of the regions r_i and r_j, yielding a time-to-delivery of

$$2b \frac{n}{s^2} \frac{a}{s} = \frac{2b\sqrt{n}\sqrt{a}}{k^{3/4}} \qquad \text{(asymptotically a.e.)} \qquad (2.7)$$

The buses serving demands (r_i,r_i) would complete their tours in less time, so (2.7) is the maximum time-to-delivery.

If $k=s(s-1)$, we are able to do better since it is then possible to have all buses busy all of the time. Again with s subregions in R, let each bus service a demand (r_i,r_j), $i\neq j$ and $1/2(s-1)$ of each of the demands (r_i,r_i)

and (r_i, r_j). Then in each of its regions, a bus will visit

$$\frac{1}{(s-1)}\frac{n}{s^2} + \frac{n}{s^2} = \frac{n}{s(s-1)} \qquad \text{points.}$$

The total distance travelled by each bus is then

$$2b\sqrt{\frac{a}{s}}\sqrt{\frac{n}{s(s-1)}}$$

$$= 2b\left[\frac{2na}{k(\sqrt{1-4k}-1)}\right]^{1/2} \qquad \text{(a.a.e.)} \qquad (2.8)$$

For very large k the values in (2.7) and (2.8) are not of too much interest. However, it is believed that the values are then 'almost optimal': they indicate the minimal time-to-delivery as a function of k, for k large.

The schemes for the multi-bus problem that we have analysed are seen in retrospect to have a certain specialization of the buses in common. For the case in which transfers are allowed, the buses are restricted to specialist <u>regions</u>. For the case in which transfers are not allowed, the buses specialize in collecting <u>demands</u> of a certain type only.

CHAPTER III

THE AVERAGE FLOW-TIME CRITERION AND THE DYNAMIC CASE

Two inadequacies of the models in Chapter II are the distance criterion, which is not really relevant to an individual passenger, and the static nature as, in actuality, we must consider an ongoing process, with demands arriving over time.

This chapter investigates these aspects of the problem by generalizing the previous results. In the dynamic case we shall consider the average flow-time criterion, so Section 3.1 investigates this criterion for the static case. Then, Section 3.2 analyses the (single-bus) dynamic problem and its relationship with the static problem. Finally, Sections 3.3 and 3.4 study the multiple-bus dynamic problem.

Once again, details of proofs are collected in an appendix. The notation and assumptions are the same as those described in Section 2.0.

3.1 The Static, Single-Bus, Average Flow-Time Problem

The criterion used in Chapter II - that of minimizing the time-to-delivery of the final passenger - is not necessarily satisfactory for an individual passenger. For,

he is interested rather in his personal wait-time and travel-time. Therefore, consider minimizing the average flow-time, i.e. the average time-to-delivery, of all passengers. This criterion will be used in the dynamic problem; the present section is needed as a link between the static and dynamic cases.

We need some additional notation. For any given problem instance of size n, let

O_n = set of n origins

D_n = set of n destinations.

Any feasible tour T defines

$t(o_i|T)$ = time at which origin o_i is visited

$t(d_i|T)$ = time at which destination d_i is visited.

Let

$$w_n(T) = \frac{1}{n} \sum_{i=1}^{n} t(o_i|T) = \text{average waiting time of } T$$

$$f_n(T) = \frac{1}{n} \sum_{i=1}^{n} t(d_i|T) = \text{average flow-time for all destinations of } T.$$

The problem is, then, given a set of n demand pairs distributed at random in the region R with area a, minimize $f_n(T)$, with respect to all feasible tours.

We shall be solving a simpler version of this problem: essentially, we discretize the tours by considering only simple tours in S, those that visit the points in R

region-by-region(1). Then, decisions are made at discrete instants of time - whenever a bus exits from a region. The choice of a tour corresponds to the choice of a strategy; at each stage that a bus exits from a region we must decide:-

(i) which region is to be visited next,

(ii) which points are to be visited in that region.

The restricted problem is

$$\text{minimize } \{f_n(T),\ T \in S\} \tag{3.1}$$

In order to provide a Beardwood-type result for problem (3.1) we first consider a problem with an easier solution:

$$\text{minimize } \{w_n(T) + f_n(T),\ T \in S_m\} \tag{3.2}$$

We draw on classical scheduling theory for our informal discussion. Recall the following result (see e.g. Conway [6]). Given a single machine and n jobs to be processed, the average flow-time is minimized if the jobs are ordered in increasing processing-time order.

Suppose that each of the m subregions has area a/m, and that n_i points are visited in a certain region. Then, the time spent there is, for large n_i, approximately

---

(1) Recall the definition in Section 2.2.2.
It might seem that the restriction to S is a stringent one. However, for the minimum distance problems in Chapter II there is no loss of generality if we restrict ourselves to S, as we are interested in ε-optimal solutions: for any optimal tour T^* there exist an M and a T_M in S_M such that

$$Y(T_M) - Y(T^*) < \varepsilon \qquad \text{a.e.}$$

for n large enough. It is believed that this observation is valid for the average flow-time problem too, but a rigorous proof promises to be awkward.

$b\sqrt{n_i}\sqrt{a/m}$, and we can take the time spent on each of the points to be $b\sqrt{a}/\sqrt{m}\sqrt{n_i}$. (Recall that the n_i points are randomly distributed in the region and n_i is large, so we can assume that the time spent between visitations is constant.) Now, to minimize the flow-time of all <u>points</u>, the scheduling result described above implies that we must visit first the points which will require least time, i.e. we must minimize $b\sqrt{a}/\sqrt{m}\sqrt{n_i}$. To do this, use the following rule.

<u>Algorithm 4</u>

Partition the region R into m subregions of equal area. At each stage, upon exiting from a region, choose as next region that with the most feasible points, and visit all feasible points there. []

This procedure corresponds precisely with the tour T_m^o given by Algorithm 2 (ties are broken arbitrarily). It minimizes the average flow-time of all <u>points</u>, i.e. origins and destinations, which is the criterion of problem (3.2).

Observe that the tour T^o has a long average waiting time, $w_n(T^o)$: passengers wait for collection while as many previous passengers as possible are delivered. Since T^o minimizes $w_n(T) + f_n(T)$ in (3.2), we might expect that T^o works well at minimizing $f_n(T)$ in problem (3.1). This is indeed the case, and T^o is asymptotically optimal in (3.1) again. This is stated and proved (in the appendix) in Theorem 4.

More precisely, in Lemma 4.1 we derive the facts that, for the tour T_m^o given by Algorithm 2,

$$\lim_{n\to\infty} \frac{w_n(T_m^o)}{\sqrt{n}} = .575b\sqrt{a} + O(1/m) \quad \text{a.e.}$$

$$\lim_{n\to\infty} \frac{f_n(T_m^o)}{\sqrt{n}} = 1.142b\sqrt{a} + O(1/m) \quad \text{a.e.}$$

Hence we have, from the considerations above,

Lemma 4.2

Let $w_n^+ + f_n^+ = \inf\{w_n(T) + f_n(T),\ T\in S\}$

Then,

$$\lim_{n\to\infty} \frac{w_n^+ + f_n^+}{\sqrt{n}} = 1.717b\sqrt{a} \quad \text{a.e.}$$

Further, given any $\varepsilon>0$, there exists an $M(\varepsilon)$ such that for $m\geq M$ and n large enough, T_m^o is ε-optimal for problem (3.2) a.e.(2) []

Theorem 4

Let $f_n^* = \inf\{f_n(T),\ T\in S\}$.

Then,

$$\lim_{n\to\infty} \frac{f_n^*}{\sqrt{n}} = 1.1417b\sqrt{a} \quad \text{a.e.}$$

Further, given any $\varepsilon>0$, there exists an $M(\varepsilon)$ such that for $m\geq M$ and n large enough, T_m^o is ε-optimal for problem (3.1) a.e. []

Thus, we have that the tour T^o is nicely robust. It minimizes a cost relevant both to the passengers and to the operator of the system.

---

(2) Recall the footnote at Corollary 2. We require this stronger statement for the proof of Theorem 4.

It is worth observing that by combining Lemma 4.2 and Theorem 4, we obtain the result that the tour T^o minimizes with respect to $T \in S$ any criterion

$$g.w_n(T) + f_n(T), \qquad g \in [0,1].$$

As before, it is possible to compare the performance of suboptimal schemes. Let us look at tours in S_m, for fixed m.

It was seen earlier that the 'Michigan Scheme' corresponds to T_1^o. The bus spends approximate time $b\sqrt{na}$ collecting the passengers and $b\sqrt{na}$ delivering them. The average time-to-delivery is then

$$\begin{aligned} f_n(T_1^o) &= b\sqrt{na} + b\sqrt{na}/2 \\ &= 1.5b\sqrt{na}. \end{aligned}$$

Comparing with f_n^*, we obtain that this tour solves the problem to within $1.5/1.1417 = 1.314$, and we lose over 30% by restricting ourselves to tours in S_1.

Nonetheless, the sequence $\{f_n(T_m^o), m=1,2,...\}$ converges quite rapidly. For m=2, $f_n(T_2^o) = 1.330b\sqrt{na}$ and so T_2^o solves the problem to within $1.330/1.1417 = 1.165$. For m=3, $f_n(T_3^o) = 1.269b\sqrt{na}$ solving the problem to within 1.112. Finally, for m=4, $f_n(T_4^o) = 1.238b\sqrt{na}$, solving the problem to within 1.08. Thus, by restricting ourselves to S_4 we lose only 8%.

3.2 The Dynamic Case: Single-Bus Problem

Possibly the most glaring inadequacy of the previous problem formulations was the assumption that all demands were known at the starting time, $t=0$. For a more realistic formulation we must assume that the demands arise as time proceeds according to some intermittent arrival process. In this section we study the extension of the results from the static case to this dynamic version.

As before, we shall assume that any particular demand has both origin and destination drawn independently from a uniform probability distribution over the region R and that the demands are all independent of one another. For simplicity, we assume[3] that demands arise at a constant rate, q.

To complete the problem specification we need a performance criterion. It is clear that, to minimize distance only without regard to time, is uninteresting: we can always wait until next year and then collect and deliver all demands together. So we choose to minimize the average flow-time. Thus, if a particular demand arises at time t_1 and the destination is visited at time t_2, the flow-time for this demand is t_2-t_1; this is averaged over all demands.

---

(3) This assumption is unnecessarily harsh, and certainly we could assume a Poisson arrival process with mean q. However, the assumption is consistent with our informal presentation and it eliminates the need for many 'expectations'.

As in Section 3.1 we discretize the problem by restricting ourselves to the class of simple tours, S.

A standard easy heuristic for solving the dynamic version of a decision problem is to use the solution of the static version recursively - this loosely corresponds to the 'open-loop-optimal feedback' control scheme described by Dreyfus [7]. For our problem the technique is the following. At each stage compute the optimal (minimal flow-time) static tour for visiting all remaining points, assuming that no new demands will arise. Use this tour to yield the initial decision for the dynamic problem - i.e. which region to visit next, and which points there. By the time this has been implemented, new demands will have arisen. With this new initial state, recompute an optimal static solution.

In general, the heuristic is suboptimal as it does not take into account that future demands will arise. But, consider its application to our problem. We claim that the static minimum flow-time problem which is faced at any particular stage is solved by the same principle given in Algorithm 4: at each stage visit the region with most feasible points and visit all such feasible points there. We do not prove this claim; a detailed proof would be as long as, and more involved than, the previous proofs, and no additional insights would be gained.

The following is a translation of the heuristic into a proposed solution for the dynamic problem - Algorithm 4 is applied recursively. We call the tour T_m^o; it is believed that no confusion with the static tour can arise.

<u>Algorithm 5</u> (yielding the tour T_m^o)

Partition the region R into m subregions of equal area. In each subregion which is visited, perform a travelling salesman tour on all feasible points. Upon exiting from a subregion, choose as the next one that with most feasible points. []

In order that the proposed static tour at each stage be (almost) optimal, the arrival rate q must be very large. Then the tour T_m^o will visit the regions consecutively. For, with probability 1, the region with the most feasible points will also be the one with longest elapsed time since it was last visited.

How well does Algorithm 5 perform? Using a fundamental result from queuing theory we argue the following lemma in the appendix.

<u>Lemma 5.1</u>

If the system is in steady-state, then with probability 1, the tour T_m^o given by Algorithm 5 minimizes, among tours in S_m, the average flow-time of all passengers. []

With this knowledge it is possible to give a Beardwood-type relationship for the dynamic problem when it has reached an equilibrium. Since q is the arrival rate of passengers, in order that the system be in a stable steady-state, we must serve q passengers per unit time. Thus, on average, 2q points must be visited per unit time. For the tour T_m^o in S_m described above, let

θ_q = the time spent in each subregion when the arrival rate is q.

During this time, $2q\theta_q$ points must be visited; so by Theorem 1 (recall that the buses travel at unit speed),

$$\lim_{q\to\infty} \frac{\theta_q}{\sqrt{2q\theta_q}} = b\sqrt{\frac{a}{m}} \quad \text{a.e.}$$

i.e.

$$\lim_{q\to\infty} \frac{\theta_q}{q} = 2b^2\frac{a}{m} \quad \text{a.e.} \tag{3.3}$$

Define $P_q = m\theta_q$ as the 'period' - this is the time required for the bus to perform a circuit around R, visiting all subregions. Then,

$$\lim_{q\to\infty} \frac{P_q}{q} = 2b^2 a \quad \text{a.e.} \tag{3.4}$$

From this it is easy to derive the formula below (see the appendix).

<u>Theorem 5</u>

Let $F_{q,m}$ be the optimal average flow-time for tours in S_m when the arrival rate of passengers is q. Then,

$$\lim_{q\to\infty} \frac{F_{q,m}}{q} = 2b^2 a(1+1/m) \quad \text{a.e.} \tag{3.5}$$

Hence, if F_q is the optimal average flow-time for tours in S, then,

$$\lim_{q\to\infty} \frac{F_q}{q} = 2b^2a \qquad \text{a.e.} \qquad (3.6)$$

[]

This theorem is analogous to Theorem 4, but for the dynamic problem. It is trivial that, by restricting ourselves to S_m instead of to S, we solve the minimum flow-time problem to within $(1+1/m)$ a.e.

3.3 Dynamic Multiple-Bus Problem: Transfers allowed

Now let us extend the last section to the k-bus problem.

To introduce the approach, consider a 'Dynamic Michigan Scheme', a tour which we call (again) $T_1^{k^o}$.

Algorithm 6.1 (yielding $T_1^{k^o}$)

Partition the region R into k subregions, each of area a/k. Let all buses meet at a common transfer point(4) at times $\theta, 2\theta, 3\theta, \ldots$. During time $[i\theta, (i+1)\theta]$ each bus visits its own subregion collecting passengers who arrived during $[(i-1)\theta, \theta]$ and delivering passengers who arrived during $[(i-2)\theta, (i-1)\theta]$. At the transfer point each passenger transfers onto the bus serving his required destination. []

(4) Recall Section 2.3.2 where this was justified by Figure 2.2. We omit any mention here of a 'line-haul' bus.

Here, $\theta = \theta_q$, the steady-state time spent by each bus in its region. During this time $q\theta_q$ additional demands arise with $2q\theta_q/k$ new points in each region. Then, as before,

$$\lim_{q\to\infty} \frac{\theta_q}{\sqrt{2q\theta_q/k}} = b\sqrt{\frac{a}{k}} \qquad \text{a.e.}$$

i.e. $$\lim_{q\to\infty} \frac{\theta_q}{q} = \frac{2b^2a}{k^2} \qquad \text{a.e.}$$

The period P_q equals θ_q, so

$$\lim_{q\to\infty} \frac{P_q}{q} = \frac{2b^2a}{k^2} \qquad \text{a.e.}$$

Now seek the average time-to-delivery. Consider a passenger arriving during time $[0,\theta_q]$ where 0 and θ_q are transfer times. This passenger is collected during $[\theta_q,2\theta_q]$ and is delivered during $[2\theta_q,3\theta_q]$. The average time-of-arrival is $\theta_q/2$, the pick-up time is on average $3\theta_q/2$ and the delivery-time is on average $5\theta_q/2$ (the expected wait-time is θ_q, as is the expected travel-time). So the average flow-time by this algorithm, viz. $F^M_{q,k}$, is $2\theta_q$ satisfying

$$\lim_{q\to\infty} \frac{F^M_{q,k}}{q} = \frac{4b^2a}{k^2} \qquad \text{a.e.} \qquad (3.7)$$

With $k=1$ this tour is precisely T^o_1 and thus (3.7) is but equation (3.5) with $m=1$, as it should be.

The tour $T^{k^o}_1$ is in S_k. A demand arising during $[0,\theta_q]$ must wait until the next period before it can be collected. This wait-time can be improved by further dividing the region of each bus (with area a/k) into m

equal subregions of area a/km each. Let each bus visit all of its m subregions sucessively, visiting as many origins as possible in each, and all the feasible destinations - now 'feasible' refers to the fact that the corresponding origin has been visited and the passenger is on the bus. After visiting their m subregions, the buses all meet at the same transfer point as before, passengers transfer, and the cycle is repeated. We call the resulting tour $T^k_{1,m}$; it is in S_{mk} and it requires a single transfer per period. Compared with $T^{k^o}_1$, the waiting-time of passengers is reduced and passengers with origin and destination both served by the same bus might not have to visit the transfer point. It can be shown (this is a special case of Theorem 6 to follow) that this flow-time, viz. $F^k_{q;1,m}$, satisfies

$$\lim_{q\to\infty} \frac{F^k_{q;1,m}}{q} = \frac{b^2qa}{k^2}\left[3 - \frac{1}{k} + \frac{1+k}{mk}\right] \quad \text{a.e.} \qquad (3.8)$$

With $k=1$, (3.8) corresponds to (3.6) as it should.

In what sense are the two schemes $T^{k^o}_1$ and $T^k_{1,m}$ optimal? The answer might be forseen from the results of the static case and Theorem 5. An optimal tour in S_{mk} would require m transfers per period, one at each time that the buses (simultaneously) complete tours in their subregions of area a/km. This would ensure that a passenger who has been collected is transferred as soon as possible onto the bus that visits his destination. As a special case, T^k_1 is optimal in S_k.

But, as pointed out earlier, it may be of interest to limit the number of transfers per period. Suppose we insist on only one transfer per period, and consider tours in S_{mk}. Then $T^{k^o}_{1,m}$ is optimal!

More generally, consider the k-bus problem in which we have at most h transfers per period, and we allow tours in S_{khm}. Then we have the following algorithm yielding the tour $T^k_{h,m}$.

<u>**Algorithm 6.2**</u> (producing $T^k_{h,m}$)

Divide the region R into khm subregions of equal area. Designate hm subregions for each of the k buses and label them $r^i_1, r^i_2, \ldots r^i_{hm}$, i=1,...k. Let bus i visit successively its regions $(r^i_1, r^i_2, \ldots r^i_m)$. Following this, all buses meet at a preassigned transfer point. Then bus i visits regions $(r^i_{m+1}, \ldots r^i_{2m})$ and a second transfer occurs. After h such transfers, the buses again visit regions $(r^i_1, \ldots r^i_m)$ and the cycle repeats.

Every time a bus visits a region it visits as many points as possible in that region; these are all the uncollected origins and all the feasible destinations of passengers already on board. []

<u>**Theorem 6**</u>

The tour $T^k_{h,m}$ is optimal (asymptotically a.e.) for the k-bus minimal flow-time problem when tours are restricted to S_{khm}, and only h transfers are allowed per

period.

Let $F^k_{q;h,m}$ be the average flow-time resulting from this tour. Then,

$$\lim_{q\to\infty} \frac{F^k_{q;h,m}}{q} = \frac{b^2a}{k^2}\left[2 + \frac{1}{mh}(1+1/k) + \frac{1}{h}(1-1/k)\right] \quad \text{a.e.} \qquad (3.9)$$

[]

From (3.6), when there is a single bus, the optimal average flow-time is $2b^2aq$. From (3.9), when there are k buses, the conceptual optimal tour in $S = S_\infty$ with continuous transfers (i.e. $h\to\infty$) has a flow-time of $2b^2aq/k^2$, and we have a k^2-fold improvement. This is analogous to the static case(5); each of the k buses visits (1/k)th of the total area a and has an effective arrival rate of q/k.

Using Equation (3.9) it is possible to investigate the optimality of suggested algorithms with respect to any restricted problem formulation.

For example, among tours in which only a single transfer is allowed per period, what is lost by using the tour T^o_1 ? Well,

$$f_q(T^o_1) = \frac{4b^2aq}{k^2} \quad \text{a.e.}$$

Letting $m\to\infty$ with $h=1$ we get from (3.9) that the optimal tour has a flow-time of $b^2aq(3-1/k)/k^2$ and so T^o_1 solves the problem to within

---

(5) In the static case in Section 2.3.2, the total distance travelled was (asymptotically) proportional to $\sqrt{n}\sqrt{a}$, yielding a k-fold improvement.

$$\frac{4}{3-1/k} = 1 + \frac{1}{3-1/k} \quad \epsilon \quad [4/3,2]$$

and we lose at least $33\frac{1}{3}\%$. This is considerably more than the loss of 6% obtained for the static problem. For the dynamic problem, then, more is to be gained by having m large.

Or, what can be gained if we allow two transfers per period instead of one? With a single transfer we have a flow-time $(m\to\infty)$ of $b^2aq(3-1/k)/k^2$. With two transfers the value is $b^2aq(5/2 - 1/2k)/k^2$. The gain is

$$1 - \frac{5/2 - 1/2k}{3 - 1/k} = \frac{1 - 1/k}{2(3 - 1/k)} \simeq 17\%$$

for k large.

These numerical values should not be taken too literally. But they do give qualitative insights into the relative optimality of suggested schemes.

3.4 Dynamic, Multiple-Bus Problem; Transfer-Free

Let us lastly consider some dynamic multiple bus schemes with no transfers. As in the static case of Section 2.3.3, we do not guarantee optimality. However, simple interesting schemes can still be studied.

3.4.1 Two-Bus Schemes

Consider the following scheme, $T'(2)$, as shown in Figure 3.1. Partition R into m subregions. Bus I visits the subregions in order $(r_1, r_2, \ldots r_m, \ldots)$, and bus II visits them in order $(r_m, r_{m-1}, \ldots r_1, \ldots)$. Bus I serves demands of the form (r_i, r_j) where $|(j-i) \bmod m| < m/2$, and bus II serves demands with $|(j-i) \bmod m| > m/2$. Demands (r_i, r_i) are shared equally between the buses.

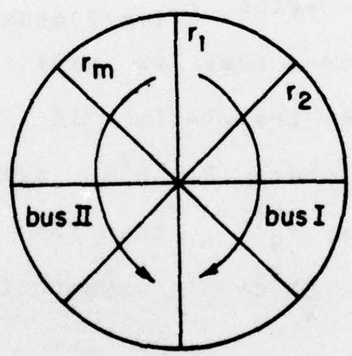


Figure 3.1 Two-bus dynamic scheme without transfer

For this scheme, let θ be the time spent by each bus in each of its subregions. During a time-interval θ, $q\theta$ demands (i.e. $2q\theta$ points) arise. In equilibrium, each bus must collect $q\theta$ points in each subregion, i.e. $\theta = b\sqrt{q\theta}\sqrt{a/m}$, and $\theta = b^2qa/m$. The time for a bus to make a circuit of the m subregions is $P = m\theta = b^2qa$. For m

large, the average wait-time is $P/2$ and the average travel-time is $P/4$. Thus, for m large, the average flow-time for $T'(2)$, viz. $F_q^{'(2)}$, is

$$F_q^{'(2)} = \frac{3}{4}b^2qa \qquad \text{asymptotically a.e.}$$

A second two-bus scheme $T''(2)$ is similar to that above, but with both buses travelling through the m subregions in the same order $(r_1, r_2, \ldots r_m, \ldots)$. Each bus always visits as many points as possible in each subregion it enters. We can suppose too that the buses are staggered, with bus II in region $r_{(i+m/2)\bmod m}$ when bus I is in region r_i. For this scheme we have that, when m is large, the average travel-time is $P/2$ and the average wait-time is $P/4$, where $P = b^2qa$ again. Thus, if the average flow-time is $F_q^{''(2)}$, then for m large,

$$F_q^{''(2)} = \frac{3}{4}b^2qa \qquad \text{asymptotically a.e.}$$

Although both schemes T' and T'' have the same average flow-time, T' may be more desirable for a number of reasons. First, for T' the average wait-time is longer and the average travel-time shorter than those of T''. So, with T', passengers have more 'coffee time at home'. Second, with T'' the buses will not always remain staggered because of additional disturbances and uncertain travel-times. In fact, the buses will tend to merge together (a similar phenomenon is described by Newell [15]) so additional controls, causing additional delays, will be

needed. Third, suppose that R is itself a 'subregion' within a larger system, and with a transfer point just before r_1, say. Then T'' is unfair in that a passenger originating in r_1 will always have a longer travel-time than one originating in r_m. The variance of the travel-times will be smaller for T'.

3.4.2 A Three-Bus Scheme

Finally, consider a three-bus scheme $T(3)$ that is a simple extension of $T'(2)$ above. The region R is divided into $3m$ subregions, each with area $a/3m$. The buses perform routes as shown in Figure 3.2; a passenger is served by the most suitable bus, that is, one that will provide a shortest travel-time for him. Let θ be the time spent by a bus in each of its subregions. During this time $2q\theta$ points arise, and for equilibrium $2q\theta/3$ points must be visited by the bus. So, $\theta = b\sqrt{2q\theta/3}\sqrt{a/3m}$ and $\theta = 2b^2qa/9m$ asymptotically a.e. Each bus has period $P = 2m\theta = 4b^2qa/9$. When m is large, the average wait-time for a passenger is $P/2$. The average travel-time is clearly less than this, and can be shown to be $P/4$. Hence the average flow-time $F_q^{(3)}$ is $3P/4$, i.e. when m is large,

$$F_q^{(3)} = \frac{b^2qa}{3} \quad \text{asymptotically a.e.}$$

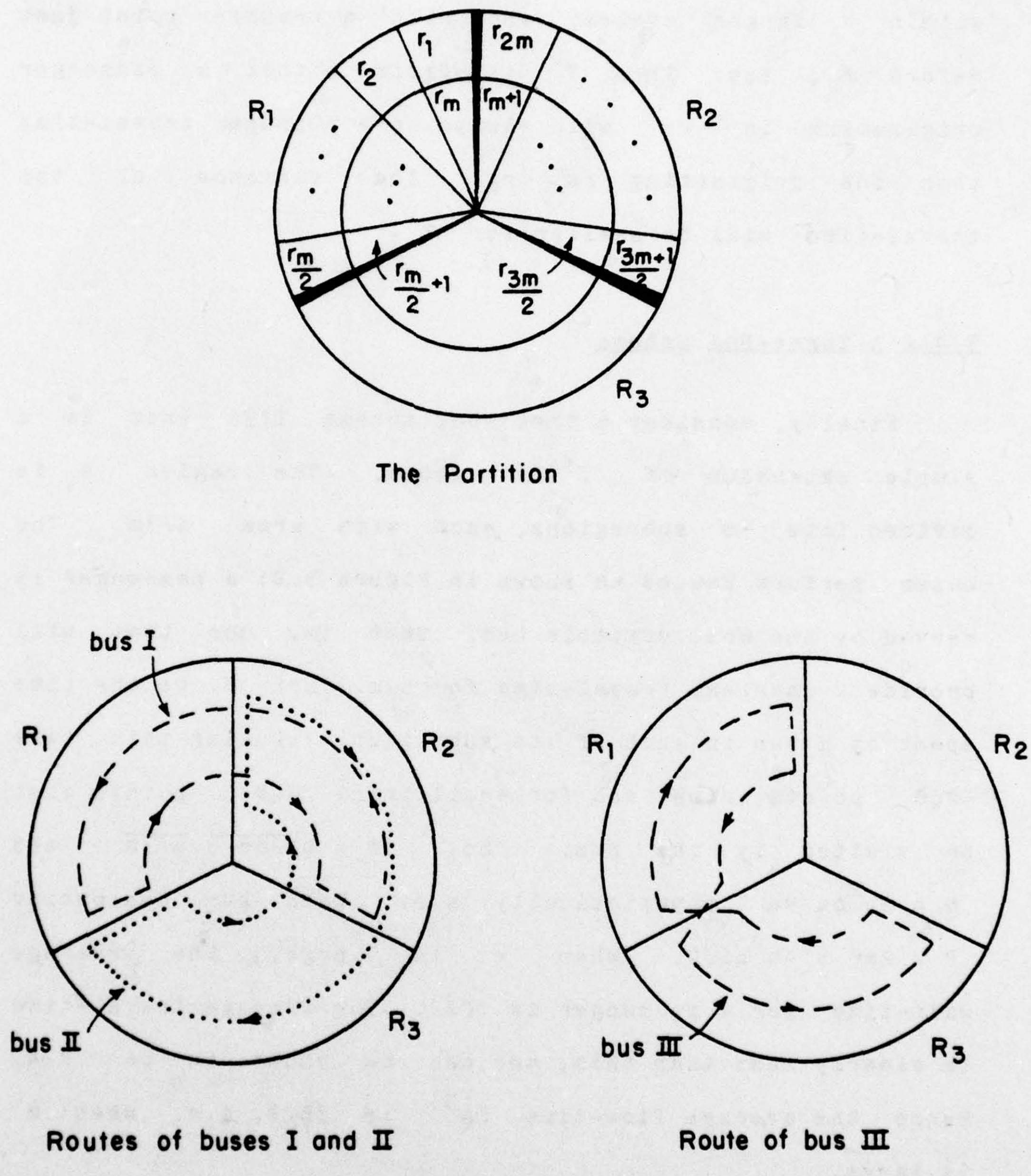


Figure 3.2 A three-bus dynamic scheme without transfers.

We conjecture that this is the minimal average flow-time for the three-bus transfer-free case. However, note again that a passenger requiring service from r_1 to r_{2m} has a very long travel-time even though the regions are adjacent. A six-bus scheme might have two buses travelling in opposite directions on each of the routes shown in Figure 3.2, and this would reduce the discrimination.

CHAPTER IV

TOWARDS THE PRACTICAL DESIGN OF DIAL-A-RIDE SYSTEMS

4.1 Introduction

In Chapters II and III we have presented a mathematical theory for idealized bus problems, which we have tacitly claimed is relevant to the real dial-a-ride problem. In order to justify this claim we must relate the real and idealized problems and draw upon the theory to obtain implications for real design. In fact, we would like to go even further and propose an approach for dial-a-ride scheduling, which is of both theoretical and practical interest. It is this that is the goal of the present chapter.

If a dial-a-ride system is required in a certain area, how should the service be designed and scheduled? Our response is essentially described by the following three planning steps. First, at the most basic level, it must be decided what approach is to be taken. The proposed approach is through the decentralization of the large problem: each bus is required to specialize only in passengers of a certain 'type', determined by their origin-destination pair. Once this approach has been adopted, the second aspect of design is that of designating the decomposition; to do this

we must specify the roles to be played by the individual buses, we must select subregions and plan transfers. At the third level is the more detailed operational design; here, we must establish decision rules that will determine the behaviour of the buses when passengers arrive.

This trinity yields a class of scheduling algorithms which we wish to propose. The class is described essentially by five 'elements' in Section 4.2. We draw upon the theory of Chapters II and III to motivate these qualitative elements and also to derive quantitative techniques for design. In Section 4.2 too, some of the theoretical assumptions and their limitations emerge more clearly. Adjustments are therefore required if the quantitative approach is to be practically meaningful, and some of these are discussed in Section 4.3. Finally, to give an improved grip on the ideas, we describe an example in Section 4.4, a 'pseudo-real situation', in which the approach is applied.

4.2 A Design Approach

This section describes, in some generality, an approach towards the design of dial-a-ride scheduling algorithms with which we shall be concerned. Given are a region where the transportation system is to be designed, and an anticipated probability distribution according to which passenger

demands will arise. A certain scheduling facility (a computer?), receiving these demands, is to determine suitable tours for a fleet of buses.

The following are five elements that characterise a class of dial-a-ride scheduling algorithms:

(i) Decomposition of Area and Specialization of Buses

(ii) Simple Tours

(iii) Rules for Visiting Subregions

(iv) Transfer Points

(v) Method of Quotation.

Our interest in these algorithms is motivated by the models of Chapters II and III. We shall discuss the five elements in detail and draw upon the theory to explain and to justify them. For this, it must be assumed that the idealized models do indeed focus upon the quintessentials of the real problem. Of course, for tractability, many assumptions and simplifications have had to be made; some of the more important of these will be discussed in Section 4.3. Despite the assumptions, the many reasonable qualitative implications do yield interesting algorithms. And the algorithms, when extended to the real problem, have many features of practical interest to boot.

Undoubtedly, fundamental to the idealized algorithms in Chapters II and III is their decomposed and discrete nature: the buses specialize in certain demand-types only, and the tours are all 'simple'. For example, recall Theorems 2 and

3 and the tour T_m^o which approximate to any required accuracy the optimal performance[(1)]. Even very simple-minded algorithms of this form yield good performance. Thus we are led to claim that the following two elements of design are worthwhile.

<u>(i) Decomposition of Area and Specialization of Buses</u>

Initially, the region in which the transportation system is to be constructed, is partitioned into a large number, m, of subregions. Then, a passenger requesting service from a subregion r_i to a subregion r_j can be directly classified according to his demand-type, (r_i,r_j) - there are m^2 such demand-types. Now, assign to each of the buses a fixed set of subregions which it visits. Each bus is to specialize within its subregions in serving certain demand-types only. That is, it visits only their origins, or destinations, or both. A bus may also periodically visit 'transfer points'. At these, it collects additional passengers and it delivers a subset of its current passengers, each of whom will be collected by some other bus.

---

(1) The measure of performance is distance travelled or average flow-time. The schemes have thus been found to be robust, minimizing criteria relevant both to the operator of the system and to the passengers. We shall discuss these, and other measures of performance, subsequently.

By the discussion in Section 2.3, we should assign buses to the demands in such a way that all buses share an equal portion of the 'load' and that each bus is busy all the time. We should also ensure that the buses are always suitably dispersed over the area: this will promise good service to the future demands that arise according to the given probability distribution.

The number m of subregions to which each bus is assigned is important, particularly in dynamic problems. However, this number does not need to be too large before good performance is attained (see Sections 3.1 and 3.2). When the demand distribution is uniform, the subregions can all simply be of equal area, yet otherwise arbitrary[(2)]. In practice, however, physical constraints such as boundaries of towns, major roads, and so on, will influence this choice.

Once the buses have specialized, the demand-type of a passenger who requests service will trivially determine the bus, or sequence of buses, to which he is to be assigned. The central scheduler that receives the demand performs no long search, but only informs the relevant bus and provides the book-keeping.

<u>(ii) Simple Tours</u>

Passengers' demands arise over time. Depending

---

(2) Additional remarks will be made in Section 4.3.

upon the actual requests, tours for the buses within their subregions must be determined. Any point which a bus visits must be 'feasible' - it must be an origin or a destination of a passenger who is already aboard. (These points are of course within the bus's speciality.) The points are to be visited in some 'optimal' order. The problem is simplified by constructing tours in discrete stages, whenever a bus enters a new subregion. At this time of entry, let the bus calculate an optimal travelling salesman tour in the upcoming subregion; any point visited there is feasible at this instant of time. The calculation is to be done on-line.

We have termed tours of this form 'simple'. Recall that the use of these tours was justified for the static problems of Chapter II, and it was claimed that they were suitable for the dynamic problems too (Chapter III, page 43, footnote (1)).

If the subregions are small and only a few points are to be visited in each, it may be easy for a driver to 'see' what route he should take. Otherwise, he might have a minicomputer aboard that finds a tour for him. Efficient approximation algorithms for the travelling salesman problem are given in [12,3]. In this way, the calculation of tours is decentralized - it reduces to a number of small discrete

problems. By choosing the number of subregions carefully, it is possible to guarantee that the total computation time is bounded by a polynomial function of q, the arrival-rate of passengers (see Karp [9]). Note too that the communication between the buses and the central scheduler does not need to be continuous, so relatively cheap communication equipment should suffice.

(iii) Rules for Visiting Subregions

Two questions remain whenever a bus exits from a subregion at a time t_i: which region should it enter next, and which points should it visit there? Well, in any region it enters, it must visit all points within its specialization that were feasible there at time t_i. And, it should visit its regions successively(3), in a fixed order.

These rules were proved valid for the case in which demands were uniformly distributed, in Section 3.1. Essentially, they require that buses 'always do as much as is possible'. They imply that a bus will not deviate in a backward direction from its prospective route to service

---

(3) This should be further qualified. It holds when the distribution of points is uniform. In the case of more general distributions, a generalized rule may be required. For example, visit next the region with largest 'point density' (i.e. the number of feasible points per unit area) - see Section 4.3(i).

newly arriving demands, as this may delay future passengers. After t_i, feasible points that arrive in the subregion are not collected: a passenger who just misses the bus must wait for the next one to enter the subregion. Even though some passengers are inconvenienced, their contribution to the average performance is kept small.

It may happen that, in a particular subregion, there are no feasible points. Then, the bus can simply bypass that subregion or travel through it as efficiently as possible to the next. We can think of the bus as travelling through its subregions along a fixed route and deviating within each subregion to collect and deliver passengers there. Passengers then know the direction in which the bus is proceeding, and should be happy, knowing roughly where they are headed.

(iv) Transfer Points

The transfer points (see (i) above) at which a bus stops from time-to-time are typically on the boundary of some subregions. They must be considered during the calculation of the optimal travelling salesman subtours. Our description has been sufficiently general to allow, for example, that some buses have 'null subregions', travelling only between transfer points; these are the familiar line-haul buses.

Transfer points were shown to be necessary in Sections 2.3.2 and 3.3 if the delay incurred can be neglected. For, they provide vast improvements in performance, and this improvement increases with the number of transfers. However, too large a number of transfer points is clearly impractical, and passengers might attach an additional cost to the number of transfers they have to make. Also, the theoretical approach of having all buses meet at the same transfer point may not be desirable:- the additional trip from the region to the transfer point will not be negligible when the number of passengers is small. In this case, the line-haul bus effectively creates the single transfer in a practical way.

(v) Quotation Times

In the statement of the dial-a-ride problem in Section 1.1 it was required that, at the time a passenger requests service, he be quoted a time of collection and delivery. The scheduling procedure we have described has not needed such a quotation. We regard it as a separate secondary problem, requiring that an estimate of the 'load' on the system, and hence an estimate of the speed of the buses, be made.

The approach which we have described guarantees a certain average service to the set of all passengers. Thus, if quotations are made to individual passengers, and if these are then introduced as additional constraints, the average service to passengers will be reduced. In practice, quotations are required, but the scheduler need not be committed to them. We believe that simple constant estimates of the times of collection and delivery will usually suffice. It is known through which subregions a bus is to travel: the time spent deviating in each subregion could be estimated and quotation times could easily be provided. So, an accurate prediction of travel-time between points is avoided.

The elements (i)-(v) above outline, in somewhat general terms, a class of scheduling algorithms. The basic approach is described by (i) and (ii) and we have claimed that this approach is justified in an asymptotic probabilistic sense. The detailed operational design is described by (iii) for a special case, and in the next section it will be seen that this can be extended fairly generally. What is lacking, then, is a methodology for selecting a suitable decomposition and for organizing the manner in which the buses specialize.

For the special-case idealized models of Chapters II and III no such methodology is required since the solution is relatively easy: when the demands are uniformly

distributed over the region, partition it into subregions of equal area. Then, in the manner of Algorithm 6.2 the buses specialize in disjoint subregions and have cyclic tours through them when the arrival-rate of passengers is large. In general, however, the answer is not as easy and difficulties, besides the nonuniformity of distributions, abound in practice. Indeed, let us remark upon a few practical considerations which indicate that a general methodology for the determination of subregions and bus-routes would surely be very complex.

First, what criteria actually determine good service? The criteria assumed in Chapters II and III, viz. the distance travelled by the buses and the average flow-time of passengers are not all-important, but are only first approximations to real requirements. There are, for example, the variance of passengers' travel-times (measuring, in a sense, the 'fairness' to the passengers) and the maximum of all passengers' travel-times (since we wish to provide all demand-types with good service). Or, there are questions of the reliability of the system (what happens if a bus breaks down?) and its stability (what effect will additional delays and uncertain travel-time have?).

Furthermore, a new dial-a-ride system would have to suit the existing circumstances. Boundaries of towns, rivers or roads may impose physical constraints on the way

in which the region can be subdivided. We may wish to supply a special quality of service to certain demand-types (e.g. to disabled persons, or see Example 4.4). A dial-a-ride system must be compatable with existing public transportation systems: it should supplement them, and yet could also depend upon them. For example, existing bus- or subway-routes might reduce the number of line-haul buses that are needed if the subregions are designed to take advantage of them.

Thus, dial-a-ride schemes must be tailor-made to suit individual complex requirements. Even though a general methodology has not been proposed, our study has yielded many qualitative insights, and no doubt additional principles could be obtained with further research.

On the other more favourable hand, however, we wish to suggest that our approach does yield a practical design tool. The emphasis throughout Chapters II and III was upon the comparison of algorithmic performance (see Definition (2.1)). In general, it is possible to quantify the performance of a suggested scheme whenever a distribution of demands, a decomposition of the region into subregions and an organization of specialist buses, are given. For, one can estimate the time spent by each bus in each of its subregions (using Beardwood's formula, Theorem 1) and evaluate its dynamic behaviour as it travels around its

subregions and visits its transfer points[4]. The service supplied to the various demand-types by the suggested scheme can then be evaluated with respect to average flow-time and to many other performance criteria. Thus, a suggested scheme can be quantitatively studied. Usually, closed-form analytic formulas should be obtainable to express its characteristics. Then, investigations can be made to analyse the behaviour of the scheme when the parameters - e.g. the number of buses, the demand distribution, the number of transfer points, and so on - are varied. Finally, alternative schemes, with alternative regional decomposition and bus-specialization can be evaluated and compared. These are analytic studies, and avoid the need for complex simulations. An example that briefly illustrates such a quantitative investigation is given in Section 4.4.

It is in order, finally, to make some additional remarks on the design approach which we have expounded in this section.

The approach can be applied on a broader scale. The 'city' may itself be a module, connected to others within a larger system. Also, the method is relevant to 'one-to-many' systems, [21], where passengers have a common origin and wish to be taken to various destinations. For, it is easy to assign incoming passengers to buses

---

(4) This was illustrated for the simple models in Chapter III.

simultaneously, merely treating the single origin as a transfer point. 'Many-to-one' systems are similar.

In [21], the existence of 'advance requests', where passengers request service some time in advance, is regarded as important. With our approach these demands are easy to handle. If the 'period' of a bus is P - this is the time it takes for the bus to make a circuit of all its subregions - the passenger can be promised collection during $[t_1, t_1+P]$ for any suitable time t_1. Then, when the region is entered, the passenger will be collected if this can be during $[t_1, t_1+P]$, otherwise the collection is deferred. Delivery by a certain time can similarly be guaranteed in advance.

There is a further important implication of the decomposition described in (i) above. The success of a transportation system depends ultimately upon the co-operation of the bus-drivers, particularly if they have decision-making capabilities. Since each bus has certain fixed demand-types assigned to it, it is possible to measure the efficiency of individual drivers. If each driver travels as well as possible through his subregions, the total system performance will be maximized. So, some incentive might be given to the drivers that will encourage them to perform well and, for example, to reduce the number of unnecessary stops. Furthermore, since the drivers communicate with the central scheduler only intermittently,

the 'big-brother' paranoia of continuously-watched drivers is reduced.

4.3 Quantitative Aspects of Design

In Section 4.2 it was described how suggested schemes can be investigated by quantifying their behaviour. Investigations such as these are limited by the many theoretical assumptions of Chapters II and III, since inaccuracies are then produced in the resulting measurements. In this section we discuss some of the more important of these assumptions and investigate to what extent they can be relaxed. We indicate how the inaccuracies might be reduced by making adjustments in the applied formulas.

4.3.1 The Demand Distribution

In deriving formulas measuring, for example, the average flow-time for a suggested scheme (see Theorems 5 and 6) Theorem 1 is used to obtain the time spent by a bus in each of its subregions. For this, it must be assumed that the points which are visited in each subregion are drawn from a uniform probability distribution there: this was the case in our previous analyses since the demand pairs were 'random'.

transportation system is to be designed, and an anticipated probability distribution according to which passenger

Suppose, now, that the points to be visited in a subregion are not uniformly distributed, but arise according to a probability density function, p, say. Then it has been shown in [2] that the optimal tour-length L_n satisfies

$$\lim_{n\to\infty} \frac{L_n}{\sqrt{n}} = b \int_R p^{1/2} dv \qquad \text{a.e.}$$

When $p = 1/a$, we have the uniform distribution. Since $\int_R p^{1/2} dv \leq \sqrt{a}$, nonuniformity of the distribution decreases L_n. Consequently, the additional structure in the nonuniform distribution can be used to advantage.

The expression above might be used to improve the analyses when the distribution is nonuniform, and when all origins and destinations are independent of one another. The discretization imposed by simple tours will remain useful, however. For, if the whole region is divided into m subregions and if we approximate the distribution p to be uniform within each subregion, then by making m large, the approximation can be made arbitrarily negligible.

Now, consider a bus visiting its subregions. What rule (i.e. element (iii)) should it use? It is not hard to see that in order to minimize average flow-time the bus should always 'do as much as possible' (compare with Lemma 5.1). Thus, whenever it enters a subregion it should always visit as many points as possible there. But, which subregion should it visit? If, in each of its subregions r_i, $i=1,\ldots,m$ there are n_i feasible points to be visited

and if then $T_i(n_i)$ will be the time spent in each, choose i so as to minimize $T_i(n_i)/n_i$. Even when the arrival-rate q of demands is large, it may be the case that the bus does not visit its subregions successively. In general, the analytic calculation of average flow-times will be more difficult than before, but will still be feasible.

Note that these remarks are not true if the origin and destination in a demand pair are correlated.

Finally, it is in order to comment upon the stochastic process by which demands arrive over time. We assumed a constant arrival-rate q, and the analysis could be extended for a Poisson process with mean q. In practice, q will change over time. This change will be relatively slow - at certain times of the day alternative designs for the decomposition and specialization might be selected, so as to suit the changing operating conditions. It is believed that a few discrete such changes should suffice in practice.

4.3.2 The Metric

We have taken as the distance between any two points in the city their euclidean separation - i.e. the length of the straight line joining them. In reality, the distance travelled by a bus is not as simple. If the actual distance can be approximated by a constant multiple of the euclidean distance, the analysis of Chapters II and III remains valid

(1) The measure of performance is distance travelled or average flow-time. The schemes have thus been found to be robust, minimizing criteria relevant both to the operator of the system and to the passengers. We shall discuss these, and other measures of performance, subsequently.

with very minor modifications.

Alternatively, the distance between two points x, y may be better approximated by a multiple of the 'Manhattan metric', or L^1, given by

$$\text{distance}(x,y) = |x_1-y_1| + |x_2-y_2|.$$

Again the results hold, this time with a different value of the constant b [2].

More general metrics that take into consideration indirect or faster roads, promise to be an order of magnitude more difficult: for, there is then no result corresponding to Theorem 1.

4.3.3 Asymptotic Approximation

Possibly the most glaring inadequacy of the approach is that it is asymptotic in the number of passengers. The expressions which are obtained are laws of large numbers, and require that the number of passengers and their arrival-rate q be large. To apply the formulas we would ideally seek a 'central limit' result allowing us to estimate the probable magnitude of the errors when q is small, and to determine the size of q needed for reliable estimates.

Sadly, a general result of this nature is not available. A theoretical analysis is hard and results depend, for example, upon the shape of the region. For R

a square, the best that it has been possible to do analytically on the convergence of Theorem 1, is to show that[5]

$$\frac{E[L_n]}{\sqrt{n}} = b\sqrt{a} + O(\log n/n)$$

where $E[L_n]$ is the expected length of an optimal travelling salesman tour for a random instance of the problem of size n. It might be hoped that the actual rate of convergence is faster than this.

In this spirit, let us investigate empirically the convergence of Theorem 1. We shall consider only points which are distributed randomly in a unit square, and shall seek to estimate $E[L_n]$ as a function of n. Comparing $E[L_n]$ with $b\sqrt{n}$ will yield an expected error in the prediction. This error will enable us to bound, in a sense, the error due to the asymptotic approximation in dynamic versions of the problem as well.

The value of $E[L_n]$ was estimated for $n=20,30,...$ as follows. Random travelling salesman problems were generated by scattering n points on the unit square (x and y coordinates were generated uniformly and independently). A number of instances of each problem size n were solved[6].

---

(5) This can be derived in the manner of Beardwood's Lemma 6.

(6) It should be remarked that a similar experiment was performed by Christofides and Eilon [4]; however, details of their numerical results and the algorithms used were not provided.

are given in [12,3]. In this way, the calculation of tours is decentralized - it reduces to a number of small discrete

The algorithm used was a heuristic of Lin and Kernighan [12]: it uses modest computation time and has been found to 'practically guarantee optimality'. For n large the computation time was still high, however, and fewer problems were sampled. The results obtained are tabulated in Figure 4.1. From these, we estimate the value of b to be about .765; Figure 4.2 compares the empirical expected tour-length with the value of $.765\sqrt{n}$. In Figure 4.3 we plot $E[L_n]/\sqrt{n}$ versus n, indicating how it converges to .765. The curve $.765 + 4/n$ is given for comparison: it seems clear that $E[L_n]/\sqrt{n}$ converges at a faster rate than $O(1/n)$.

It is interesting to note that for every random problem-instance tried, with $b = .765$,

$$b\sqrt{n} \le L_n.$$

So, with high probability the tour-length predicted by Theorem 1 is <u>less</u> than the actual value. Indeed, from Figure 4.3,

$$b\sqrt{n} \le E[L_n] \le b\sqrt{n} + 4/n \qquad (4.1)$$

This yields a fairly close bound on the magnitude of the error implicit in Theorem 1.

Now let us obtain an estimate of the convergence rate for a bus problem - we are interested primarily in the dynamic version. In evaluating the performance of a scheme the time spent by a bus in each of its subregions is estimated, so we can utilize the results above. We consider

For example, visit next the region with largest 'point density' (i.e. the number of feasible points per unit area) - see Section 4.3(i).

Problem size n	10	20	30	40	50	60	70	80	90	100	150	200
	3217	4084	4488	5388	6032	6506	6680	7067	7632	8086	9419	10841
	2965	3536	5018	5246	5706	5952	6468	6879	7179	7612		10924
	2561	4350	4270	5697	6121	6223	6610	6899	7675	7979		
	3205	3569	4822	5119	5659	5902	6308	6740	7401	8128		
Optimal	2855	4359	4578	5218	5833	6218	6391	7095	7607	7839		
tour-lengths	2901	4371	4315	4926	6118	5871						
L_n	3239	3499	4161	5155	5429	6107						
for random	2865	4116	4642	5398	6137	6193						
instances	3269	4109	5193	4594	5787	6336						
x 1000	3558	3839	4972	5383	5877	5966						
	2354	3762	4215	5280	5651	6497						
	2813	3710	4779	5174	6188	5899						
	3000·	3864	4711	5305	5375	6349						
	2568	3782	4607	5390	5749	6132						
	2654	3815	4588	5023	5299	6401						
$E[L_n]$	2935	3917	4624	5220	5797	6170	6491	6936	7499	7929	9419	10883
$E[L_n]/\sqrt{n}$	.928	.876	.844	.825	.820	.797	.776	.775	.790	.793	.769	.7695

Figure 4.1 Random instances of the travelling salesman problem, generated on the unit square.

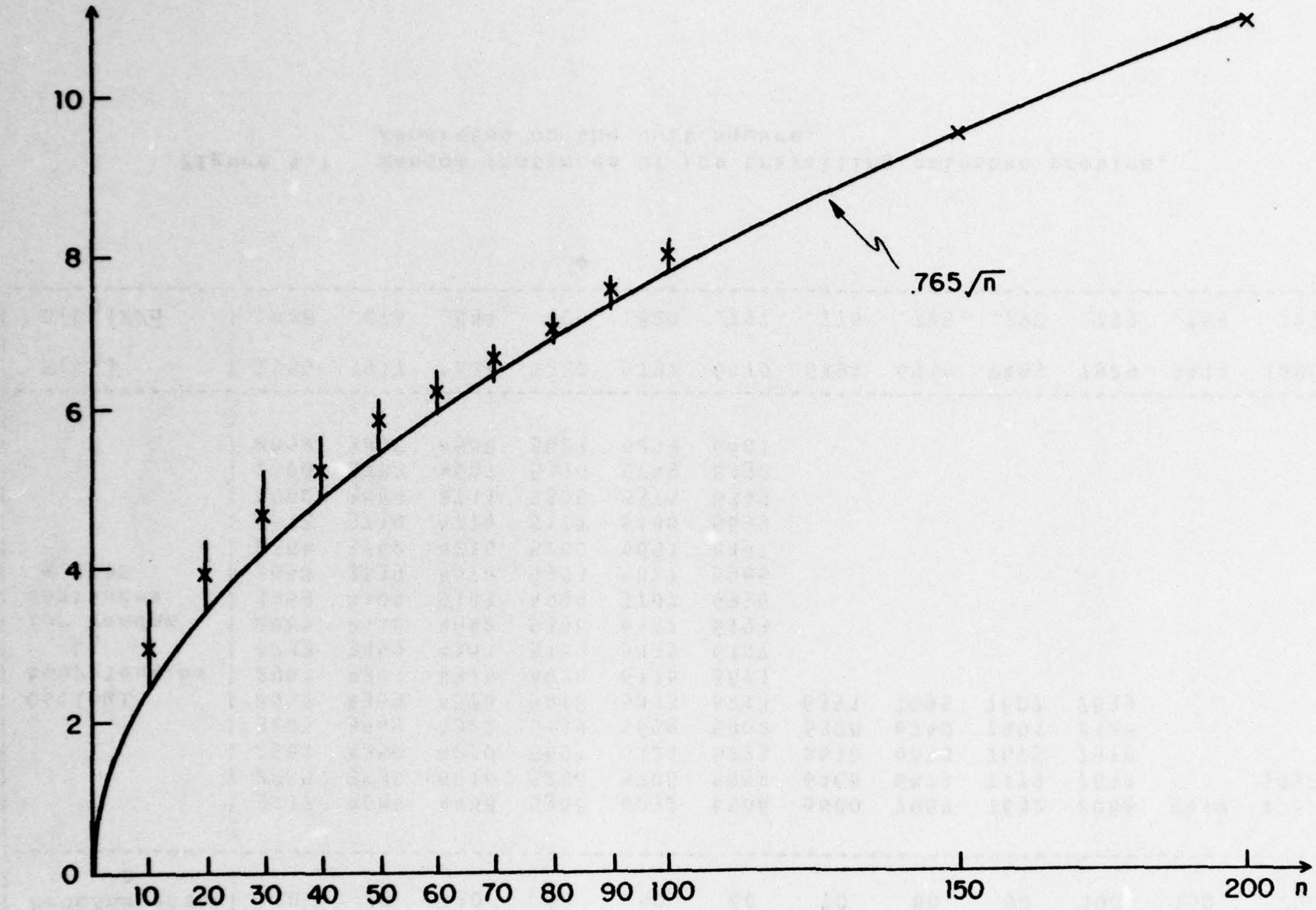


Figure 4.2 Asymptotic convergence to Beardwood's formula. For each n, the range of experimental results is shown; $E[L_n]$ is denoted by x. For comparison, the curve $.765\sqrt{n}$ is plotted.

familiar line-haul buses.

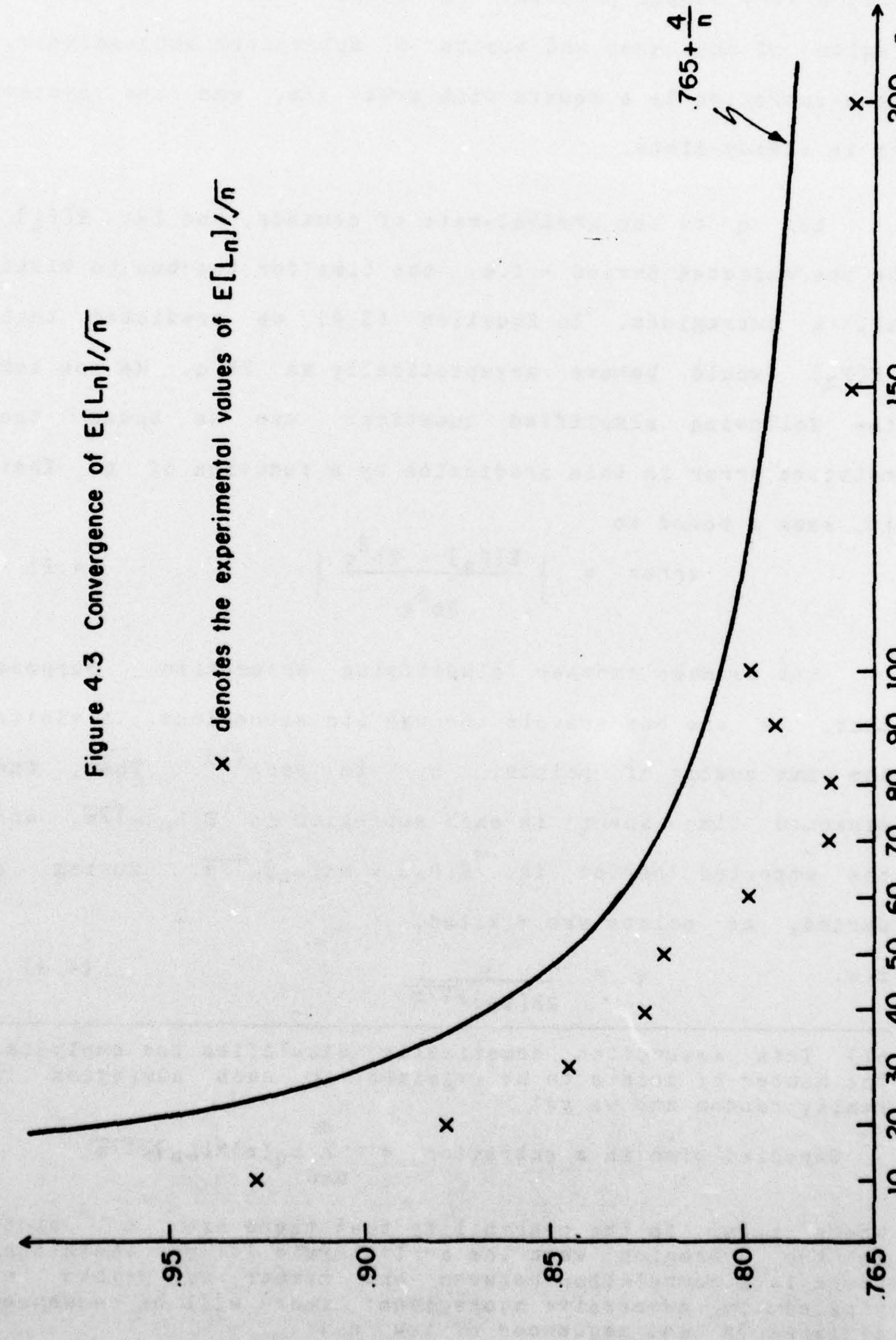


Figure 4.3 Convergence of $E[L_n]/\sqrt{n}$

only a very simple problem: a single bus is serving a region of unit area and visits m subregions successively. Each subregion is a square with area $1/m$, and the system is in steady-state.

Let q be the arrival-rate of demands, and let $E[P_q]$ be the expected period - i.e. the time for the bus to visit all m subregions. In Equation (3.4) we predicted that $E[P_q]$ would behave asymptotically as $2b^2q$. We now ask the following simplified question: can we bound the relative error in this prediction by a function of q. That is, seek a bound to

$$\text{error} = \left| \frac{E[P_q] - 2b^2q}{2b^2q} \right| \tag{4.2}$$

Let us make another simplifying assumption. Suppose that, as the bus travels through its subregions, it visits the same number of points, n, in each[7]. Then, the expected time spent in each subregion is $E[L_n]\sqrt{1/m}$, and the expected period is $E[P_q] = mE[L_n]\sqrt{1/m}$. During a period, mn points are visited,

i.e.
$$q = \frac{n}{2E[L_n]\sqrt{1/m}} \tag{4.3}$$

---

(7) This assumption drastically simplifies the analysis. The number of points to be visited in each subregion is really random and we get

$$\text{Expected time in a subregion} = \sum_{n=0}^{\infty} p_q(n)E[L_n]\sqrt{1/m}$$

where $p_q(n)$ is the probability that there are n points in the subregion when the arrival-rate is q. (Note that there is a correlation between the number of points n visited in successive subregions: there will be sequences of large n and sequences of low n.)

rivers or roads may impose physical constraints on the way

From (4.2),

$$\text{error} = \left| \frac{E[L_n] - b^2 n/E[L_n]}{b^2 n/E[L_n]} \right|$$

Using (4.1),

$$\text{error} \leq \left| \frac{(b\sqrt{n} + 4/n)^2 - b^2 n}{b^2 n} \right|$$

$$= \left| \frac{8b\sqrt{n} + 16/n}{b^2 n^2} \right|$$

With $b = .765$, the relative error given by (4.2) is less than 12% when n, the number of points visited in each subregion, is 20; it is less than 7% when n is 30; and it is less than 5% when n is 35. *For any n,* (4.3) determines the corresponding arrival-rate, q.

4.3.4 Additional Uncertainties

The following are a few additional issues (mainly related to stochastic variations) that we have neglected and that we now wish to note.

(i) The speed of a bus is not constant as has been assumed. The travel-time between two points is uncertain and is subject to traffic and road conditions.

(ii) Delays that occur when a stop is made have been neglected. A bus may have to wait for a passenger to appear and for his impedimenta to be loaded. Similar delays occur for alighting passengers.

(iii) A bus is also detained at transfer points. The time

taken is not instantaneous, but is a function of the number of passengers who transfer. Indeed, if buses are to meet at transfer points as was required in Chapters II and III, waits are inevitable.

(iv) From a passenger's viewpoint, the time spent on a line-haul bus (including his wait) may be appreciable. In our asymptotic analysis these, as well as the time taken for a bus to travel a fixed bounded distance were neglected.

(v) Consider a bus travelling a circuit around its subregions. In Chapter III we calculated the equilibrium value of the average flow-time. If, now, the bus is delayed, additional passengers still arrive, and the time spent in the following subregions will be longer than anticipated. So, the delay will persist and the equilibrium will only be re-achieved gradually.

Because of these assumptions, there will be yet further errors in the derived formulas. However, it is believed that many of them could be taken into account in a practical analysis. For example, studies of urban traffic behaviour might be used to more fully understand the travel-times; and we might be able to estimate the time spent by passengers on the line-haul system. Then, the tools which we have proposed could be further sharpened and calibrated.

(4) This was illustrated for the simple models in Chapter III.

Finally, a more sophisticated technique for the detailed operational control - one that will take into consideration the random fluctuations - can be suggested. Consider specifying in advance a schedule of times for the buses to be at transfer points. Then, if a bus is running late, it can improve its speed by selecting carefully which of its passengers it should collect: some passengers remain uncollected until the bus (or the next suitable bus) re-enters their subregion. This feature would require additional machinery if it is to be analysed. A measure of the 'state' of each bus would have to be defined, and it would be nontrivial to determine the schedule in advance - certain 'slack' would be required to allow for random disturbances. The schedule would have to be 'stable' with respect to reasonable perturbations and 'flexible', so that changes can easily be made if things go awry. To study these topics would require new methods of some 'stochastic scheduling theory'.

4.4 An Example

Let us suppose that we have a city, R, in which we want to construct a dial-a-ride transportation system. This city is essentially composed of two distinct regions as shown in Figure 4.4 - downtown, R_d, with area a_d, and suburbs, R_s, with total area a_s. There is already an efficient public transportation system operating in R_d.

of unnecessary stops. Furthermore, since the drivers communicate with the central scheduler only intermittently,

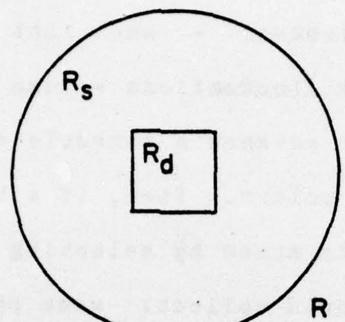


Figure 4.4 An example : the city

For most of the day there is heavy traffic travelling between R_s and R_d in both directions. The dial-a-ride scheme is to be designed for this scenario.

Suppose we can approximate the distribution of demands that are to be served by our system by the following. Any particular demand pair (o,d) is of the form (R_s,R_d) with probability p/2, $p \in (0,1)$; it is of the form (R_d,R_s) with probability p/2; and it is of the form (R_s,R_s) with probability (1-p). This distribution, then, is nonuniform. The system is not required to serve demands (R_d,R_d). Further, the value of p is large, i.e. fairly close to unity.

As remarked, the approach which we wish to adopt is one in which R is decomposed into subregions and in which buses specialize. The theory of Chapters II and III has not

'random'.

given a general methodological procedure for designing this decomposition when demands are nonuniform, but the following could well be reasonable for the present example.

In order to discourage persons who travel between R_s and R_d from using their private vehicles, let us provide them with attractive transportation, say a transfer-free service. We describe first a scheme for only these (R_s,R_d) and (R_d,R_s) demands. Consider a division of R_s into k regions, $r_1,\ldots r_k$, each of area a_s/k, for an integer k. It might be desirable to have natural divisions here, described by boundaries of suburbs, roads and so on. This is depicted in Figure 4.5.

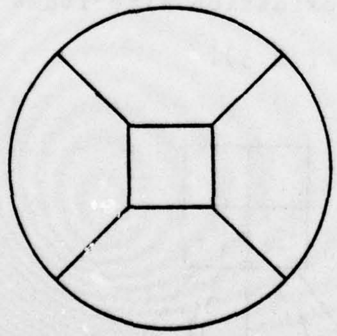


Figure 4.5 Partition of suburbs, R_s

Now allocate a bus b_i to the region r_i, for $i=1,\ldots k$; bus b_i will serve passengers with origin or destination in r_i and will travel also in R_d. Each bus

AD-A046 162 HARVARD UNIV CAMBRIDGE MASS DIV OF ENGINEERING AND --ETC F/G 12/2
SCHEDULING DIAL-A-RIDE TRANSPORTATION SYSTEMS: AN ASYMPTOTIC AP--ETC(U)
SEP 77 D M STEIN N00014-75-C-0648
UNCLASSIFIED TR-670 NL

2 OF 2
AD
A046162

END
DATE
FILMED
12-77
DDC

will travel a 'simple tour' (see Section 2.2.2), so subdivide r_i, $i=1,...k$ and R_d further, into m subregions each. Bus b_i travels a route as shown in Figure 4.6. It makes a circuit around its $2m$ subregions, visiting them successively. If it enters a subregion $r_{i,j}$ at a time t, it visits there all possible points that are feasible at t. These are the origins of demands that have arisen during the time since $r_{i,j}$ was last entered, and all destinations in $r_{i,j}$ of passengers who are already aboard b_i. The tour in $r_{i,j}$ is an optimal travelling salesman tour, and can be calculated at t. If the number of points to be visited in $r_{i,j}$ is small, the driver might easily be able to choose his best tour. Otherwise, there might be a minicomputer aboard that calculates the tour for him. Efficient approximation algorithms that might be used for this are those in [12,3].

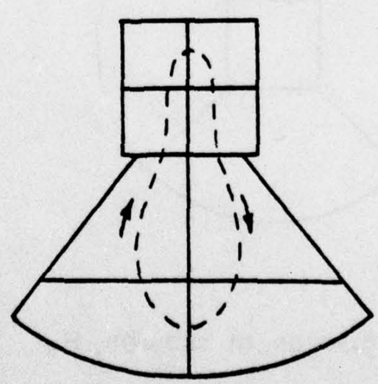


Figure 4.6 Route taken by a single bus in a region

The buses communicate with a central computer (or scheduler) that informs them of the points to be visited in their subregions. This computer monitors all demands and assigns them to the buses; the assignment decisions are trivial - no long or hard search is needed.

With k regions of R_s and k buses, each travelling at unit speed, the asymptotic equilibrium average flow-time for the (R_d, R_s) and (R_s, R_d) demands in the above scheme, can be calculated using the techniques developed in Chapter III as follows. Consider the route of a particular bus. Let θ be its period, the time for it to make a circuit of its 2m subregions[8]. Let θ_d be the time spent in each subregion of R_d, and let θ_s be the time spent in each subregion of r_i. Then,

$$\theta = m\theta_s + m\theta_d \tag{4.5}$$

During θ, there arrive $2pq\theta/k$ points that are to be served by this bus, i.e. $pq\theta/km$ in each subregion. Thus, by Theorem 1,

$$\theta_d = b\sqrt{\frac{pq\theta}{k}}\sqrt{\frac{a_d}{m}} \tag{4.6}$$

$$\theta_s = b\sqrt{\frac{pq\theta}{k}}\sqrt{\frac{a_s}{km}} \tag{4.7}$$

From (4.5), (4.6) and (4.7),

$$\theta = \frac{b^2pq}{k}\left[\sqrt{a_d} + \sqrt{\frac{a_s}{k}}\right]^2$$

As in the proof of Theorem 5 we get that the average

---

(8) Note that we are being rather loose here. These are the equilibrium values of θ when q is large; the equations to follow hold asymptotically, with probability 1.

flow-time of these passengers is

$$F_q \; - \; \frac{b^2pq}{k}\left[\sqrt{a_d} + \sqrt{\frac{a_s}{k}}\right]^2(1+1/2m) \qquad (4.8)$$

for q large. This approximation to the value of F_q should not be taken too literally: some of its limitations are discussed in Section 4.3.

Various other similar schemes might also be suggested for serving the (R_s,R_d) and (R_d,R_s) demands. Suppose that the number of buses is increased. Then, the average flow-time will be minimized if we increase also the number of regions of R_s, and retain a single bus in each. However, it might also be desirable to have more than one bus in each region r_i. For example, two buses travelling through the subregions as shown in Figure 4.7 would provide a more equitable service to passengers (i.e. a [illegible]r variance of travel-time) and a lower value of the ma[illegible]um flow-time. This scheme is also more reliable: if a bus breaks down, its passengers can still be collected by the other bus. These are all considerations that are important in practice.

Having described how passengers travelling between R_d and R_s might be served, we now turn to suburban demands, of the form (R_s,R_s). In general we have seen that a service with transfers will yield an average flow-time that is better than that with a transfer-free service. So, consider a scheme with transfers.

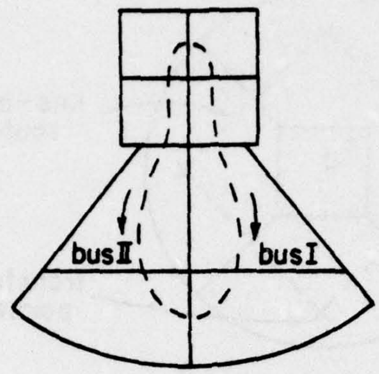


Figure 4.7 Route taken by a pair of buses in a region

With j buses, partition R_s into j regions $r_1,...r_j$ as was done before, and assign a bus to each. Also specify a transfer point in each region and a route for a line-haul bus, travelling between these transfer points, in the manner of the 'Michigan Scheme', as described in Section 1.1. This line-haul bus will play the role of a transfer point: the distance which it travels is asymptotically negligible, and it is clearly not always practical to have all regional buses meeting at the same transfer point as was required in Algorithm 3. For our example, this is sketched in Figure 4.8.

Once more, subdivide the regions r_i; the regional buses perform simple tours in their subregions in the same manner as those before, visiting the transfer point each period. The previous remarks on the number of buses in each

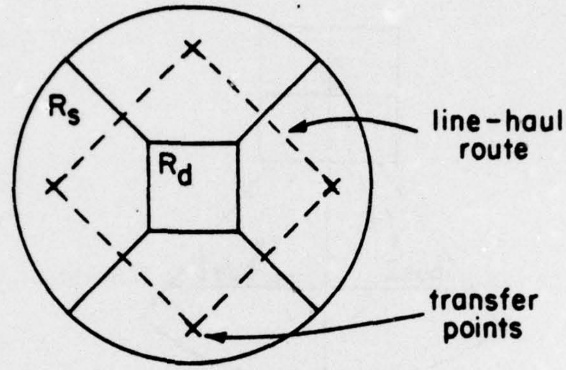


Figure 4.8 Decomposition for suburban demands

region are again pertinent.

If there are j regional buses and j regions, with m subregions each, the asymptotic average flow-time can be calculated from Theorem 6 (set k=j; h=1) to be

$$F_q \sim \frac{b^2(1-p)qa_s}{j^2}\left[3 + \frac{1}{m} - \frac{1}{j}(1-1/m)\right] + Q_1 \qquad (4.9)$$

Here, Q_1 is the average time spent by a passenger on a fixed-route bus - it depends for example upon the number of these buses and upon the actual route taken.

This scheme corresponds to a single transfer per period. It is possible to obtain further improvement in average flow-time by having, say, two transfers per period. Then, we need 2j regions of R_s, with each bus assigned to two of them, and a transfer point in each. A bus visits its two regions alternatingly, performing a tour through the subregions in each, and visiting the transfer point

thereafter (in the manner of the tour $T^j_{2,m}$ of Algorithm 6). The asymptotic average flow-time can be calculated as before to be

$$F_q = \frac{b^2(1-p)qa_s}{j^2}\left[2.5 + \frac{1}{2m} - \frac{1}{2j}(1-1/m)\right] + Q_2 \qquad (4.9a)$$

With this approach, then, it is possible to calculate the asymptotic flow-time and other performance criteria for various suggested schemes as in (4.8) and (4.9). It is indicated in Section 4.3 how adjustments in these formulas might be made to reduce their error. Then, a parametric study can be made to analyse the behaviour of alternative system designs. For example, given a fixed number of buses, *determine how they should be allocated between the suburban and central subsystems.* Or, calculate the possible gain in flow-time if the number of buses is increased or if there are more transfers on the suburban subsystem. Different schemes with alternative decomposition of the area might also be studied, and other performance criteria such as the average waiting-time or maximum travel-time might be investigated. This can all be done analytically, avoiding the need for complex simulations.

CHAPTER V

CONCLUSION

Our first aim in this report has been towards the development of dial-a-ride scheduling algorithms. We have attempted an analysis of the problem at its most fundamental level, i.e. to question what design approach should be taken. In our study we have viewed the system macroscopically - by considering the statistics of the arriving demands, we have analysed probabilistically the combinatorial aspects of the problem for suitable mathematical models. This has led to a derivation of a class of algorithms (or, schemes) for which performance can easily be evaluated, and hence to an approach towards the design of dial-a-ride systems.

The proposed approach, then, is justifiable in a mathematical sense. The justification is with respect to simple criteria that are pertinent both to the operator of the system and to the passengers. The mathematical sense is asymptotic and probabilistic, and will be discussed shortly.

From a practical viewpoint, the design approach has many attractive features. Most important is that the schemes have a decomposed, modular nature. This yields a remarkable simplicity at many levels. Computational

requirements are kept low: no long searches are needed, and tours for the buses are calculated at discrete instants of time in a decentralized fashion. The systems are simple and easy to visualize: passengers should understand their basic operation and anticipate the type of service to be provided. It is possible too to provide incentives for the bus-drivers: if each driver performs as well as he is able, good performance for the whole system will result.

A further feature is that a tool for design has been provided. It is possible to calculate analytically the dynamic behaviour of each bus, and hence to measure the performance of a suggested scheme in terms of the parameters (e.g. the number of buses, the size of the region, the number of transfer points, and so on). One can thus choose parametric values 'optimally' and can evaluate and compare alternative schemes.

As remarked, we have concerned ourselves with the fundamentals of algorithmic design. Therefore, we have had to avoid operational details - there are very many practical aspects which we have barely, if at all, touched upon. So, there is scope for much additional work. Further analytic work could focus upon a specific methodology for choosing a decomposition of the region and a specialization of the buses in subregions, particularly when the demand is nonuniform. We have but indicated (in Chapter IV) how the design tool might be sharpened and calibrated if stochastic

disturbances, low demand rates and additional travel-times are considered. A computer simulation might be used to develop these methods. A final validation of the approach would be in a real implementation: the schemes do appear promising since a related system at Ann Arbor, Michigan, has been successful.

The approach has been justified via the mathematical models of Chapters II and III. This brings us to the second concern of our report: Scheduling Theory as studied in Operations Research.

At the start, we wished to question the asymptotic approach to combinatorial optimization suggested by Karp [9]. Karp's approach has permitted a global view of the problem: when a new demand arises there is no need for a large combinatorial manipulation of the existing solution, and a small adjustment suffices. It yields simple, decentralized algorithms that are efficient (in terms of computational requirements) and yet are nontrivial. Optimality can be essentially guaranteed when the problem size is very large. When the problem size is small, the theory is no longer valid, but the resulting algorithms are nonetheless relevant. Thus, we are led to claim that Karp's approach does appear to have practical worth.

In applying [9] to our transportation problem we have generalized a result of Beardwood et. al. [2]. It is tempting to conjecture that similar generalizations,

yielding interesting practical algorithms, might be possible for other travelling salesman-like problems that abound in Operations Research.

Traditionally, scheduling has been regarded as a combinatorial problem. This emphasis has produced largely intractable problem formulations and, except for some special cases (see Conway et. al. [6]), few insights. Our study has essentially avoided combinatorics. It is believed that many of the features we have obtained are more generally relevant to scheduling applications. Particularly when there are stochastic disturbances (as is the case in most large real problems) it is felt that combinatorial aspects can be dissipated in a fairly simple way, by decomposing and discretizing the problem suitably. Then, other effects predominate: these are problems of coordination and of designing the schedule in advance. The schedule should be 'stable' with respect to reasonable disturbances and should be 'flexible' so that changes can easily be made when required. Phenomena such as these are not well understood within the present framework of Scheduling Theory.

APPENDIX TO CHAPTER II

This appendix is a collection of detailed proofs of the assertions in Chapter II. All notation and terminology is as described there. For the most part the results are proved rigorously; for the sake of clarity in Proposition 3.2 we do revert to a more informal presentation.

First we make some preliminary notational remarks, and restate the law of large numbers and Beardwood's result for completeness.

Preliminary Remarks

Throughout the appendix we shall be dealing with random variables. We must explicitly identify the underlying probability space and describe our notation.

Let $w = \{(o_1,d_1),(o_2,d_2),\ldots\}$ be a countable sequence of pairs of points in the planar region R. Let $w_n = \{(o_1,d_1),\ldots(o_n,d_n)\}$ be the first n pairs of w. The set W of all sequences w is the sample space. If we regard w as a sequence of 'random' pairs, each element of $\{o_1,d_1,o_2,d_2,o_3,d_3,\ldots\}$ being independently and uniformly drawn from the points in R, then we have a probability space. Random variables are all defined on this space. The

set of 'problem instances' of size n is $\{w_n, w \in W\}$.

For example, $Y_n(w)$ is the length of the optimal bus tour through the first n points of $w \in W$, and $Y_n(\cdot)$ is a random variable. If $X_n(\cdot)$, $n=1,2,\ldots$ are random variables then $\limsup_{n\to\infty} X_n(\cdot)$ and $\liminf_{n\to\infty} X_n(\cdot)$ are well defined random variables. If X and Y are random variables, $X \leq Y$ means $X(w) \leq Y(w)$ for all $w \in W$; $X \leq Y$ almost everywhere (a.e.) means $\Pr[X(w) \leq Y(w)] = 1$.

For clarity, in what follows, we shall omit explicit mention of the sample space and the parameterization on the w's.

Kolmogorov's Strong Law of Large Numbers [14]

Let $\{X_i, i=1,2,\ldots\}$ be a sequence of independent identically distributed random variables, such that $E(X_1) < \infty$. Then

$$\frac{1}{n} \sum_{i=1}^{n} X_i \quad \text{converges almost everywhere}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i = E(X_1) \qquad \text{a.e.}$$

2.1 The Travelling Salesman Problem

Theorem 1 (Beardwood et. al. [2])

Let R be a planar Lebesgue set with area $a>0$ and let L_n be the length of the shortest path through n points which are uniformly and independently distributed in R. There exists an absolute constant b, independent of the problem instance and of the shape of R such that

$$\lim_{n\to\infty} \frac{L_n}{\sqrt{n}} = b\sqrt{a} \qquad \text{almost everywhere (a.e.)}$$

Let us prove an easy initial lemma that will be required in many proofs to follow: if the number of points in a certain subregion is random, we can replace this random number by its asymptotic value in Beardwood's formula.

Lemma 1.0

Consider a particular region of area a in which there are p_n points; p_n is a random variable satisfying for some constant g,

$$\lim_{n\to\infty} \frac{p_n}{n} = g \qquad \text{a.e.} \tag{A1.0}$$

Let h_n be the length of an optimal travelling salesman tour on these p_n points. Then,

$$\lim_{n\to\infty} \frac{h_n}{\sqrt{n}} = b\sqrt{g}\sqrt{a} \qquad \text{a.e.}$$

Proof

By Theorem 1, we have

$$\lim_{p_n\to\infty} \frac{h_n}{\sqrt{p_n}} = b\sqrt{a} \qquad \text{a.e.} \tag{A1.1}$$

By (A1.0),

$$\lim_{n\to\infty} \sqrt{\frac{p_n}{n}} = \sqrt{g} \qquad \text{a.e.} \qquad (A1.2)$$

Let any $\varepsilon > 0$ be given. Let $\varepsilon_1 = \max\{\varepsilon/3, \varepsilon/3b\sqrt{a}\}$ and $\varepsilon_2 = \max\{\varepsilon/3, \varepsilon/3\sqrt{g}\}$. From (A1.2) we have that there exists N_1 such that for $n \geq N_1$,

$$(\sqrt{g}-\varepsilon_1)\sqrt{n} \leq \sqrt{p_n} \leq (\sqrt{g}+\varepsilon_1)\sqrt{n} \qquad \text{a.e.}$$

From (A1.1), there exists N_2 such that for $n \geq N_2$,

$$(b\sqrt{a}-\varepsilon_2)\sqrt{p_n} < h_n < (b\sqrt{a}+\varepsilon_2)\sqrt{p_n} \qquad \text{a.e.}$$

Thus, for $n \geq \max\{N_1, N_2\}$,

$$(b\sqrt{a}-\varepsilon_2)(\sqrt{g}-\varepsilon_1) < \frac{h_n}{\sqrt{n}} < (b\sqrt{a}+\varepsilon_2)(\sqrt{g}+\varepsilon_1) \qquad \text{a.e.}$$

i.e. $$b\sqrt{a}\sqrt{g} - \varepsilon < \frac{h_n}{\sqrt{n}} < b\sqrt{a}\sqrt{g} + \varepsilon \qquad \text{a.e.}$$ []

and the result follows.

2.2 Static Single-Bus Problem

Observation 2.0

The single-bus static problem is NP-complete.

Proof (See Aho [1] for the notation.)

It is easy to see that the bus problem is in NP.

Given an instance of the TSP of size n, we can transform it in polynomial time into an instance of the bus problem of size n as follows. Let the given n points be origins $\{o_1, \ldots o_n\}$, and let all have as common destination a point p, a distance of at least $D\Delta$ from every origin, where D is a very large real number and Δ is the

$\max\{\|o_i - o_j\|,\ i,j=1,\ldots n\}$. For D large enough, an optimal bus tour will visit all origins first, and only then visit p. Clearly, the subtour visiting all origins is an optimal open travelling salesman tour. [ ]

For the proof of Lemmas 2.2 and 2.3 we require an easy preliminary lemma.

Lemma 2.1

$$\frac{1}{m^{3/2}} \sum_{i=1}^{2m} \sqrt{i} = \frac{4\sqrt{2}}{3} + O(1/m)$$

Proof

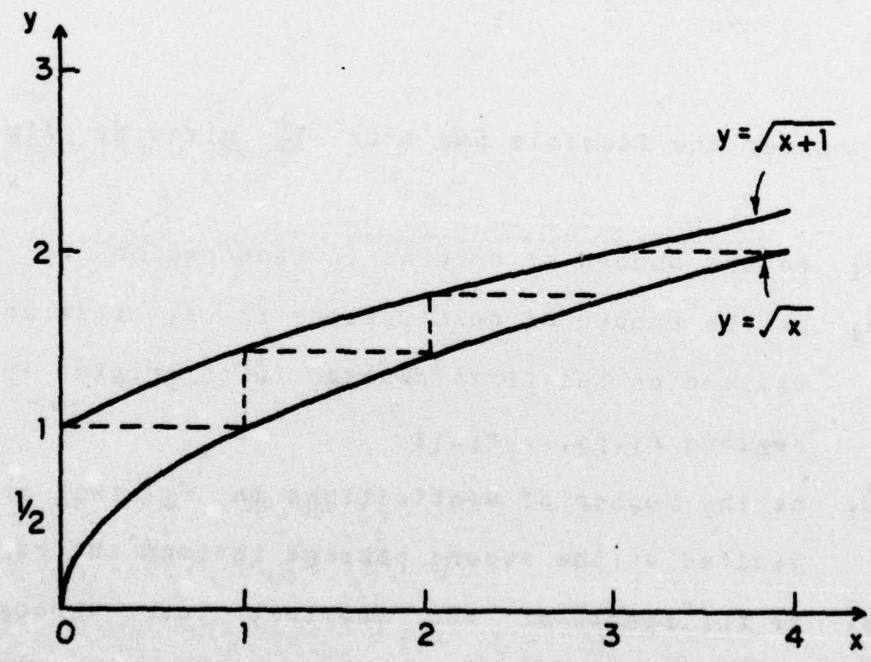


Figure A2.1 Diagram for Lemma 2.1

From Figure A2.1,

$$\int_0^{2m} \sqrt{x}\,dx \;\le\; \sum_{i=1}^{2m} \sqrt{i} \;\le\; \int_0^{2m} \sqrt{x+1}\,dx$$

i.e. $$\frac{1}{m^{3/2}}\left[\frac{2}{3}(2m)^{3/2}\right] \;\le\; \frac{1}{m^{3/2}} \sum_{i=1}^{2m} \sqrt{i} \;\le\; \frac{1}{m^{3/2}}\left[\frac{2}{3}(2m+1)^{3/2}-\frac{2}{3}\right]$$

i.e. $$\frac{2}{3}\,2^{3/2} \;\le\; \frac{1}{m^{3/2}} \sum_{i=1}^{2m} \sqrt{i} \;\le\; \frac{2}{3}(2+1/m)^{3/2} - \frac{2}{3m^{3/2}}$$

i.e. $$\frac{4\sqrt{2}}{3} \;\le\; \frac{1}{m^{3/2}} \sum_{i=1}^{2m} \sqrt{i} \;\le\; \frac{4\sqrt{2}}{3} + O(1/m)$$ []

__Lemma 2.2__

$$\limsup_{n\to\infty} \frac{Y_n}{\sqrt{n}} \;\le\; \frac{4}{3}\sqrt{2}\,b\sqrt{a} \qquad \text{a.e.}$$

__Proof__

Consider the feasible bus tour T_m^o given by Algorithm 2.

Let $\bar{o}_i$ be the number of origins in each region r_i

$\bar{d}_i$ be the number of destinations in r_i that are visited on the first passage (with origins in regions $r_1, r_2, \ldots r_{i-1}$)

$\bar{\bar{d}}_i$ be the number of destinations in r_i that are visited on the second passage through the regions.

Let s_i^1 be the length of the shortest tour through the $\bar{o}_i+\bar{d}_i$ points in r_i visited on the first passage, and s_i^2 be the length of the shortest tour through the $\bar{\bar{d}}_i$ points visited on the second passage - these are random variables.

If Y_n is the length of the optimal bus tour then

$$Y_n \leq \sum_{i=1}^{m} s_i^1 + \sum_{i=1}^{m} s_i^2 + 2m\Delta(R) \qquad \text{(A2.1)}$$

where $\Delta(R)$ is the diameter of R (so $\Delta(R)/\sqrt{n} \to 0$ as $n\to\infty$).

Now, the law of large numbers yields, as $n\to\infty$,

$$\frac{\bar{o}_i}{n} \to \frac{1}{m} \quad \text{a.e.}$$

$$\frac{\bar{d}_i}{n} \to \frac{(i-1)}{m^2} \quad \text{a.e.}$$

and

$$\frac{\bar{\bar{d}}_i}{n} \to \frac{m-i+1}{m^2} \quad \text{a.e.}$$

Thus, by Theorem 1 and Lemma 1.0,

$$\lim_{n\to\infty} \frac{s_i^1 + s_i^2}{\sqrt{n}} = b\sqrt{\frac{1}{m} + \frac{(i-1)}{m^2}}\sqrt{\frac{a}{m}} + b\sqrt{\frac{m-i+1}{m^2}}\sqrt{\frac{a}{m}} \quad \text{a.e.}$$

$$= \frac{b\sqrt{a}}{m^{3/2}}\left(\sqrt{m-1+i} + \sqrt{m-i+1}\right) \quad \text{a.e.}$$

Hence,

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}}\left[\sum_{i=1}^{m} s_i^1 + \sum_{i=1}^{m} s_i^2\right] = \frac{b\sqrt{a}}{m^{3/2}}\left[\sum_{i=1}^{2m}\sqrt{i} + \sqrt{m} - \sqrt{2m}\right] \quad \text{a.e.}$$

Inequality (A2.1) now yields

$$\limsup_{n\to\infty} \frac{Y_n}{\sqrt{n}} \leq \frac{b\sqrt{a}}{m^{3/2}}\left(\sum_{i=1}^{2m}\sqrt{i}\right) + \frac{b(1-\sqrt{2})}{m}\sqrt{a}$$

This holds for arbitrary m, so we can let $m\to\infty$, and we obtain by Lemma 2.1,

$$\limsup_{n\to\infty} \frac{Y_n}{\sqrt{n}} \leq \frac{4}{3}\sqrt{2}\, b\sqrt{a} \quad \text{a.e.}$$

[]

Lemma 2.3

$$\liminf_{n\to\infty} \frac{Y_n}{\sqrt{n}} \geq \frac{4}{3}\sqrt{2}\,b\sqrt{a} \qquad \text{a.e.}$$

Proof

First, some notation. In any problem instance, let O_n be the set of n origins and D_n be the set of n destinations. Let T_n be any optimal tour (which has length Y_n). By this we mean that T_n is that connected subset of R which is traversed in visiting $O_n \cup D_n$. A tour segment through P (where $P \subset T_n$) is the minimal connected subset of T_n that contains P.

The optimal tour T_n induces a feasible ordering on the points of $O_n \cup D_n$. Let us label the origins in O_n such that

$i<j \iff o_i$ proceeds o_j on the tour T_n.

Let d_i be the destination corresponding to o_i.

Let m be any integer, $m>1$. We partition R into m Lebesgue subregions, each of area a/m.

Let $(T_n)^1$ be the tour segment through $\{o_1, o_2, \ldots o_{[n/m]}\}$, where $[n/m]$ is the smallest integer larger than n/m. (The segment $(T_n)^1$ possibly visits some destinations en-route.) Similarly, let $(T_n)^j$ be the tour segment through $\{o_{[(j-1)n/m]+1}, \ldots o_{[jn/m]}\}$. There exist partitions of R into m Lebesgue subregions $\hat{R}_1, \ldots \hat{R}_m$ each with area a/m and bounded perimeter (length of boundary) such that the origins in $\hat{R}_j$ are precisely those in $(T_n)^j$, for $j=1,\ldots m$. Let $A_1, \ldots A_m$ form a partition

of R drawn at random from the space of such partitions[1]. We can assume that the origins in A_j are uniformly distributed there.

Let us define for each subregion A_j, $j=1,2,...m$,

$\bar{o}_{nj}$ = number of origins in A_j = $[n/m]$ or $[n/m]+1$

$\bar{d}_{nj}$ = number of destinations in A_j

$\bar{d}'_{nj}$ = number of destinations in A_j which also belong to $\{d_1, d_2, ... d_{[nj/m]}\}$ - i.e. destinations in A_j whose corresponding origins are in $A_1 \cup A_2 \cup ... \cup A_j$.

---

(1) An example of suitable A_j, $j=1,...m$, is provided by the following.

Let B_i, $i=1,...n$, be a ball around o_i with radius $\bar{\varepsilon}<1/n^2$ and with $B_i \cap B_j = \emptyset$ for $i \neq j$. That is,

$$\bar{\varepsilon} < \min\{1/n^2, \min\{\|o_i - o_j\|/2, i,j=1,...n, i\neq j\}\}$$

and $B_i = R \cap \{x, \|x-o_i\|<\bar{\varepsilon}\}$, $i=1,...n$.

Now partition $R-\bigcup_{i=1}^{m} B_i$ randomly into m equal regions X_j, $j=1,...m$, such that each X_j is a Lebesgue set with perimeter less than a fixed, suitably large number, K.

Finally, let A_j be X_j together with the balls around the origins in $(T_n)^j$, i.e.

$$A_j = X_j \cup B_{[(j-1)n/m]+1} \cup ... \cup B_{[jn/m]}$$

The perimeter of A_j is less than

$$K + n.2\pi/n^2 = K + 2\pi/n$$

(This is a bounded function of n, as is to be required in (A2.6a).)

$\bar{d}''_{nj}$ = number of destinations in $A_j \cap (T_n)^j$ - i.e. destinations which are visited by the segment $(T_n)^j$ in A_j.

Note that $\bar{d}''_{nj} \leq \bar{d}'_{nj}$.

Since the regions A_j were randomly selected, and the destinations were uniformly distributed over R, the random variables $\bar{d}_{nj}$ and $\bar{d}'_{nj}$ satisfy, by the law of large numbers,

$$\frac{\bar{d}_{nj}}{n} \rightarrow \frac{1}{m} \quad \text{a.e.} \quad \text{as } n\rightarrow\infty$$

$$\frac{\bar{d}'_{nj}}{n} \rightarrow \frac{j}{m^2} \quad \text{a.e.} \quad \text{as } n\rightarrow\infty$$

Also,
$$\frac{\bar{o}_{nj}}{n} \rightarrow \frac{1}{m} \quad \text{as } n\rightarrow\infty$$

Thus, given any $\varepsilon>0$, there exists $N_1>0$ such that, for $n\geq N_1$,

$$-\varepsilon < \frac{\bar{d}_{nj}}{n} - \frac{1}{m} < \varepsilon \quad \text{a.e.} \qquad \text{(A2.2)}$$

$$-\varepsilon < \frac{\bar{d}'_{nj}}{n} - \frac{j}{m^2} < \varepsilon \quad \text{a.e.} \qquad \text{(A2.3)}$$

and
$$-\varepsilon < \frac{\bar{o}_{nj}}{n} - \frac{1}{m} < \varepsilon \qquad \text{(A2.4)}$$

Now, the tour within A_j, viz. $T_n \cap A_j$, consists of one or more connected pieces. Discard any such piece which does not contain a point of $O_n \cup D_n$. The closure of any surviving piece will be called a j-piece. The set of j-pieces can be partitioned into two parts (see Figure A2.2):

$j^{(1)}$-pieces are those that belong to $(T_n)^j$

$j^{(2)}$-pieces are the remainder.

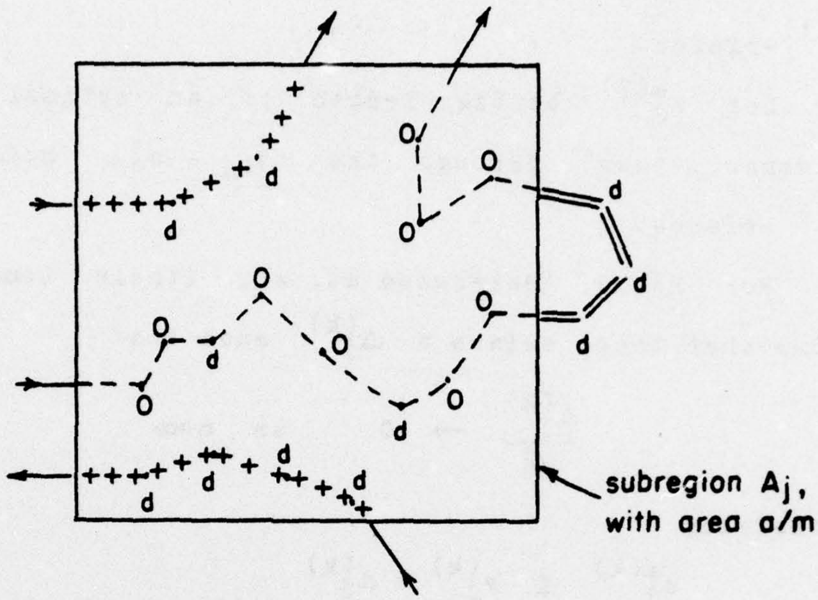


Figure A2.2 The Tour T_n^* — (i)

Part of the optimal tour T_n^* is shown. The arrows indicate the *direction traversed.*

$j^{(1)}$-pieces are shown - - - - - - -

$j^{(2)}$-pieces are shown + + + + + +

$(T_n)^j$ - segment is the ---- and ═══ path.

Thus, the $j^{(1)}$-pieces contain all $\bar{o}_{nj}$ origins in A_j as well as the $\bar{d}''_{nj}$ destinations. The $j^{(2)}$-pieces contain $\bar{d}_{nj} - \bar{d}''_{nj}$ destinations.

Let $s_j^{(1)}$ be the length of the $j^{(1)}$-pieces

$s_j^{(2)}$ be the length of the $j^{(2)}$-pieces.

Clearly,

$$Y_n \geq \sum_{j=1}^{m} (s_j^{(1)} + s_j^{(2)}) \tag{A2.5}$$

Let $s_j^{*(1)}$ be the length of an optimal travelling salesman tour through the $\bar{o}_{nj} + \bar{d}''_{nj}$ points of the

$j^{(1)}$-pieces.

Let $s_j^{*(2)}$ be the length of an optimal travelling salesman tour through the $\bar{d}_{nj} - \bar{d}''_{nj}$ points of the $j^{(2)}$-pieces.

For $k=1,2$, Beardwood et. al. (their Lemma 2) have shown that there exists a $\Delta_j^{(k)}$ such that

$$\frac{\Delta_j^{(k)}}{\sqrt{n}} \to 0 \quad \text{as } n\to\infty \tag{A2.6a}$$

satisfying

$$s_j^{*(k)} \leq s_j^{(k)} + \Delta_j^{(k)} \tag{A2.6}$$

This is obtained by constructing a feasible travelling salesman tour through the points of the $j^{(k)}$-pieces, each traversed just once, together with a part of the double circuit of the boundary of A_j. (An example is diagrammed in Figure A2.3.)

Now, there exists an N_2 such that, for all $n \geq N_2$,

$$0 < \frac{\Delta_j^{(k)}}{\sqrt{n}} < \frac{\varepsilon}{m} \quad \text{for } k=1,2;\ j=1,2,\ldots m. \tag{A2.7}$$

By Theorem 1, for $n>N_2$,

$$b\sqrt{\frac{a}{m}} - \frac{\varepsilon}{m} \leq \frac{s_j^{*(1)}}{(\bar{o}_{nj} + \bar{d}''_{nj})^{1/2}} \quad \text{a.e.} \tag{A2.8}$$

and

$$b\sqrt{\frac{a}{m}} - \frac{\varepsilon}{m} \leq \frac{s_j^{*(2)}}{(\bar{d}_{nj} - \bar{d}''_{nj})^{1/2}} \quad \text{a.e.} \tag{A2.9}$$

From (A2.8), (A2.9), (A2.2) and (A2.4), we get for $n>N = \max\{N_1,N_2\}$,

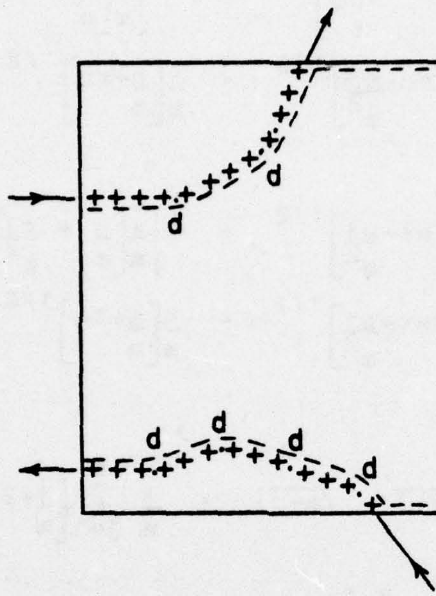


Figure A2.3 The Tour $T_n^* - (ii)$

$j^{(2)}$-pieces are shown ++++++

The tour ----- is an upper bound to the length $s_j^{*(2)}$, an optimal travelling salesman tour through the destinations in the $j^{(2)}$-pieces.

$$
\begin{aligned}
s_j^{*(1)} + s_j^{*(2)} \geq\ & b\sqrt{\frac{a}{m}}\left[\bar{o}_{nj} + \bar{d}''_{nj}\right]^{1/2} + b\sqrt{\frac{a}{m}}\left[\bar{d}_{nj} - \bar{d}''_{nj}\right]^{1/2} \\
& - \frac{\varepsilon}{m}\left[\bar{o}_{nj} + \bar{d}''_{nj}\right]^{1/2} - \frac{\varepsilon}{m}\left[\bar{d}_{nj} - \bar{d}''_{nj}\right]^{1/2} \quad \text{a.e.} \\
\geq\ & b\sqrt{\frac{a}{m}}\left[\frac{n}{m} - n\varepsilon + \bar{d}''_{nj}\right]^{1/2} + b\sqrt{\frac{a}{m}}\left[\frac{n}{m} - n\varepsilon - \bar{d}''_{nj}\right]^{1/2} \\
& - \frac{\varepsilon}{m}\left[\frac{n}{m} + 2n\varepsilon + \frac{n_j}{m^2}\right]^{1/2} - \frac{\varepsilon}{m}\left[\frac{n}{m} + \varepsilon n\right]^{1/2} \quad \text{a.e.}
\end{aligned}
$$

It is true in general that, for any $a>b>y>0$,

$$\operatorname*{argmin}_{x \in [0,y]} \{\sqrt{a+x} + \sqrt{b-x}\} = y$$

So, since $\bar{d}''_{nj} \in [0, \bar{d}'_{nj}]$, and with (A2.3) we have

$$s_j^{*(1)} + s_j^{*(2)} \geq b\sqrt{\frac{a}{m}}\left[\frac{n}{m} - n\varepsilon + \bar{d}'_{nj}\right]^{1/2} + b\sqrt{\frac{a}{m}}\left[\frac{n}{m} - n\varepsilon - \bar{d}'_{nj}\right]^{1/2}$$

$$- \frac{\varepsilon}{m}\left[\frac{n}{m} + 2n\varepsilon + \frac{nj}{m^2}\right]^{1/2} - \frac{\varepsilon}{m}\left[\frac{n}{m} + \varepsilon n\right]^{1/2} \quad \text{a.e.} \tag{A2.9a}$$

$$\geq b\sqrt{\frac{a}{m}}\left[\frac{n}{m} - 2n\varepsilon + \frac{nj}{m^2}\right]^{1/2} + b\sqrt{\frac{a}{m}}\left[\frac{n}{m} + \frac{nj}{m^2}\right]^{1/2}$$

$$- \frac{\varepsilon}{m}\left[\frac{n}{m} + 2n\varepsilon + \frac{nj}{m^2}\right]^{1/2} - \frac{\varepsilon}{m}\left[\frac{n}{m} + \varepsilon n\right]^{1/2} \quad \text{a.e.}$$

From (A2.5),

$$\frac{Y_n}{\sqrt{n}} \geq b\sqrt{\frac{a}{m}}\,\frac{1}{m^2}\sum_{j=1}^{m}\left(\sqrt{m+j} + \sqrt{m-j}\right) - \frac{\varepsilon}{m}\sum_{j=1}^{m}\left[\frac{1}{m} + 2\varepsilon + \frac{j}{m^2}\right]^{1/2}$$

$$- \varepsilon\left[\frac{1}{m} + \varepsilon\right]^{1/2} \quad \text{a.e.}$$

But, ε is arbitrary, and the coefficients of ε are bounded above, so

$$\liminf_{n\to\infty} \frac{Y_n}{\sqrt{n}} \geq \frac{b\sqrt{a}}{m^{3/2}}\sum_{j=1}^{m}\left(\sqrt{m+j} + \sqrt{m-j}\right) \quad \text{a.e.}$$

$$= b\sqrt{a}\left[\frac{1}{m^{3/2}}\sum_{j=1}^{2m}\sqrt{j} - \frac{1}{m}\right]$$

This holds for any m. Thus, by Lemma 2.1,

$$\liminf_{n\to\infty} \frac{Y_n}{\sqrt{n}} \geq \frac{4}{3}\sqrt{2}\,b\sqrt{a} \quad \text{a.e.}$$

[]

2.3 Static Multiple-Bus Problems

Lemma 3.1

$$\lim_{n\to\infty} \frac{Z_n^k}{\sqrt{n}} = c\sqrt{a} \qquad \text{a.e.}$$

Proof

Define the random variables $\bar{c}^k$ and $\underline{c}^k$ as follows.

$$\bar{c}^k = \limsup_{n\to\infty} \frac{Z_n^k}{\sqrt{n}}$$

$$\underline{c}^k = \liminf_{n\to\infty} \frac{Z_n^k}{\sqrt{n}}$$

Note first that it is feasible to have $k-1$ buses idle and to use only one. Thus,

$$Z_n^k \le Y_n$$

(Y_n is as in Theorem 2.)

Hence, $\bar{c}^k \le c$ a.e. by Theorem 2.

We now show that $c \le \underline{c}^k$ a.e.

With the fact that $\underline{c}^k \le \bar{c}^k$, we shall have the required result.

Define for $i=0,1,2,\ldots$ auxilliary problems as follows.

'Problem i': Let $Y_n^k(i)$ be the minimal total distance travelled by k buses through n random demand pairs, allowing **at most** i transfers of passengers - $Y_n^k(i)$ is a random variable.

Let
$$\limsup_{n\to\infty} \frac{Y_n^k(i)}{\sqrt{n}} = \bar{c}^k(i)$$

$$\liminf_{n\to\infty} \frac{Y_n^k(i)}{\sqrt{n}} = \underline{c}^k(i)$$

Now, for $i=0,1,2,\ldots$, we have

$$Y_n^k(i) \geq Y_n^k(i+1)$$

and

$$Y_n^k(i) \geq Y_n^k.$$

So,

$$\underline{c}^k(i) \geq \underline{c}^k(i+1)$$

and also

$$\underline{c}^k(i) \geq \underline{c}^k \qquad \text{for all } i.$$

Thus, the $\{\underline{c}^k(i), i=0,1,2,\ldots\}$ form a monotonically decreasing sequence in i, bounded below by $\underline{c}^k$. Thus,

$$\lim_{i\to\infty} \underline{c}^k(i) \overset{\Delta}{=} \bar{\Phi}^k \quad \text{exists, with} \quad \bar{\Phi}^k \geq \underline{c}^k.$$

It is shown in the Lemma 3.1.1 below that

$$\bar{\Phi}^k = \underline{c}^k \tag{A2.10}$$

For any fixed i, consider the optimal tour-length $Y_n^k(i)$ - a random variable. It is possible to justify - see Lemma 3.1.2 below - that

$$Y_n^k(i) + (i+1)(k-1)\Delta \geq Y_n^1 \tag{A2.11}$$

where Y_n^1 is the optimal distance travelled by a single bus, and Δ is the diameter of the region R.

Dividing (A2.11) by $\sqrt{n}$ and letting $n\to\infty$, we obtain by Theorem 2 that

$$\underline{c}^k(i) \geq c \qquad \text{a.e.}$$

This holds for any i; thus,

$$\bar{\Phi}^k \geq c \qquad \text{a.e.}$$

and, by (A2.10),

$$\underline{c}^k \geq c \qquad \text{a.e.}$$

This proves the result, modulo the two lemmas.

Lemma 3.1.1

We need to show that $\bar{\Phi}^k = \underline{c}^k$. Suppose to the contrary that $\bar{\Phi}^k > \underline{c}^k$ at some $w \in W$. Let $\varepsilon = \bar{\Phi}^k - \underline{c}^k > 0$. Now, there exists an I such that for $i \geq I$,

$$\bar{\Phi}^k - \varepsilon/6 < \underline{c}^k(i)$$

Also, since

$$\underline{c}^k(i) = \liminf_{n\to\infty} \frac{Y_n^k(i)}{\sqrt{n}},$$

there exists N such that for all $n \geq N$,

$$\frac{Y_n^k(i)}{\sqrt{n}} \geq \underline{c}^k(i) - \frac{\varepsilon}{6}$$

- see Royden [20, page 37].

Thus, for $i \geq I$ and $n \geq N$,

$$\frac{Y_n^k(i)}{\sqrt{n}} \geq \bar{\Phi}^k - \frac{\varepsilon}{3}$$

$$= \underline{c}^k + \frac{2}{3}\varepsilon \tag{A2.12}$$

Since

$$\liminf_{n\to\infty} \frac{Z_n^k}{\sqrt{n}} = \underline{c}^k,$$

there exists an $h > N$ such that

$$\frac{Z_h^k}{\sqrt{h}} < \underline{c}^k + \frac{\varepsilon}{3} \tag{A2.13}$$

- again, see Royden [20, page 37].

But Z_h^k is the length of an optimal tour on h demands. This tour can have only a finite number of transfers, so there exists an integer $p_h(\varepsilon)$ such that for almost every problem instance, Z_h^k requires at most $p_h(\varepsilon)$ transfers. Thus, for $i > p_h(\varepsilon)$,

$$Z_h^k = Y_h^k(i) \tag{A2.14}$$

Hence, for $n \geq N$ and $i \geq \max\{I, p_h(\varepsilon)\}$, we have from (A2.12) - (A2.14),

$$\underline{c}^k + \frac{2}{3}\varepsilon < \underline{c}^k + \frac{\varepsilon}{3}$$

This is clearly absurd and the lemma is proved.

Lemma 3.1.2

We assume that the buses start their tours from certain fixed (but arbitrary) points in R, and that all buses terminate their tours at a common final point in R. Consider the following acyclic directed graph, G, representing the tour with length $Y_n^k(i)$ for any problem instance.

Represent by nodes the k starting points, the transfer points (there are at most i of them) and the final terminal point. The tour of a particular bus i is represented by a path through G: edges into a node represent the buses which transfer passengers there, and the indegree equals the outdegree at all nodes representing transfers. An edge thus represents the path through R taken by the corresponding bus, and G represents the given set of tours with total length $Y_n^k(i)$.

An example best illustrates this definition of G. The graph in Figure A2.4 represents a tour with 7 buses and 6 transfer points.

Suppose the k tours are executed by a single bus in the following way. Pass along the same routes, 'backtracking in a straight line from each node to a previous one as far as is necessary, so that all incoming

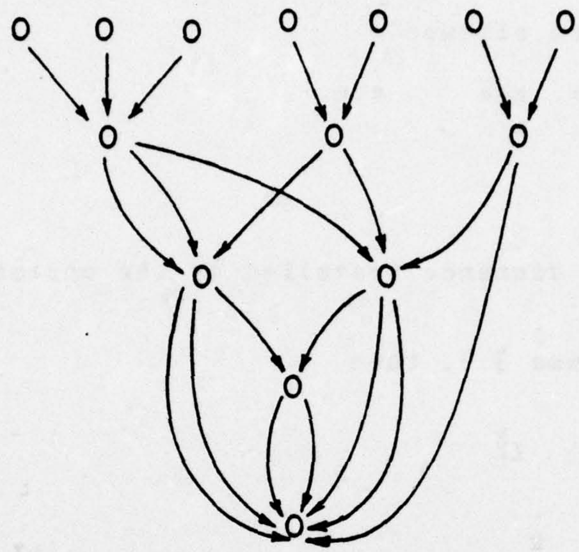


Figure A2.4 Representation of tours and transfers by a graph

edges have been executed'. (We do not feel that it is necessary to be more specific here. The idea is precisely the reverse of the standard depth-first-search algorithm - see e.g. Aho et. al. [1].) At worst we will have to backtrack k-1 times at each of i+1 nodes (the transfer points and the terminal point), and each backtrack will involve a distance of at most Δ, the diameter of R.

This leads to a feasible single-bus tour with length

$$s \leq Y_n^k(i) + (k-1)(i+1)\Delta$$

yielding inequality (A2.11).

Lemma 3.1 is thus proven. []

Theorem 3

If transfers are allowed

$$\lim_{n\to\infty} \frac{Y_n^k}{\sqrt{n}} = \frac{c\sqrt{a}}{k} \qquad \text{a.e.}$$

Proof

$\sum_{i=1}^{k} x_n^i$ is the total distance travelled by the optimal tour.

If Z_n^k is as in Lemma 3.1, then

$$\sum_{i=1}^{k} x_n^i \geq Z_n^k$$

Now, $\qquad kY_n^k \geq \sum_{i=1}^{k} x_n^i \qquad$ by definition of Y_n^k.

Dividing by $\sqrt{n}$ and letting $n\to\infty$, we obtain, by Lemma 3.1,

$$\liminf_{n\to\infty} \frac{kY_n^k}{\sqrt{n}} \geq c\sqrt{a} \qquad \text{a.e.} \tag{A2.15}$$

Consider the tour $T_m^{k^o}$ given by Algorithm 3. For this tour let

$s_n^i(m)$ be the total distance travelled by bus i

$h_n^i(m)$ be the total length of the travelling salesman tours covered by bus i in its regions.

Then,

$$s_n^i(m) \leq h_n^i(m) + m\Delta$$

where Δ is the diameter of the region R.

Now, by Lemmas 1.0 and 2.2,

$$\limsup_{n\to\infty} \frac{h_n^i(m)}{\sqrt{n}} \leq mc\sqrt{\frac{1}{m}}\sqrt{\frac{a}{m}} + O(1/m) \quad \text{a.e.}$$

Hence,

$$\limsup_{n\to\infty} \frac{s_n^i(m)}{\sqrt{n}} \leq c\sqrt{a} + O(1/m) \qquad \text{a.e.}$$

By the optimality of Y_n^k we have

$$Y_n^k \le \max\{s_n^i(m),\ i=1,2,\ldots k\} \qquad \text{for any } m.$$

Thus,

$$\limsup_{n\to\infty} \frac{Y_n^k}{\sqrt{n}} \le c\sqrt{a} + O(1/m) \quad \text{a.e. for any } m.$$

Letting $m\to\infty$,

$$\limsup_{n\to\infty} \frac{Y_n^k}{\sqrt{n}} \le c\sqrt{a} \quad \text{a.e.} \tag{A2.16}$$

The inequalities (A2.15) and (A2.16) prove the lemma.

[]

<u>Proposition 3.2</u>

The 3-bus fixed-route scheme, described in Section 2.3.3 yields a time-to-delivery Y_n^{3F} satisfying

$$\lim_{n\to\infty} \frac{Y_n^{3F}}{\sqrt{n}} = \frac{1}{3\sqrt{3}}(\sqrt{3} + \sqrt{2} + 1)b\sqrt{a} \quad \text{a.e.}$$

$$= .798b\sqrt{a} \quad \text{a.e.}$$

<u>Proof</u>

We use the same notation as that used for describing the algorithm in the text. Our derivation will be informal, but hopefully transparent. We need only consider bus I since the other buses have analogous tours and so distances travelled will be the same (asymptotically a.e.).

Bus I first visits regions $(r_1, r_2, \ldots r_m)$. In each r_i

$[(m-i+1)+m]n/9m^2$ origins are collected and

$[i-1]n/9m^2$ destinations are delivered.

So, by Lemma 1.0, there is a tour-length of $b\sqrt{a/3m}\sqrt{2n/9m}$ in r_i.

Second, the bus visits $(r_{2m}, r_{2m-1}, \dots r_{m+1})$. In each r_i

$[m+(i-1)]n/9m^2$ origins are collected and

$[(m-i)+m]n/9m^2$ destinations are delivered.

The tour-length in each of these m regions is $b\sqrt{a/3m}\sqrt{(3m-1)/9m^2}\sqrt{n}$.

Finally, the bus visits $(r_1, \dots r_m)$ again. In each r_i

$(m+1)n/9m^2$ destinations are delivered.

So in r_i there is a tour-length of $b\sqrt{a/3m}\sqrt{(m+1)/9m^2}\sqrt{n}$.

The total distance travelled by the bus is thus

$$\frac{b\sqrt{a}\sqrt{n}}{3\sqrt{3}}(\sqrt{2} + \sqrt{3-1/m} + \sqrt{1+1/m})$$

$$= \frac{b\sqrt{a}\sqrt{n}}{3\sqrt{3}}(\sqrt{2} + \sqrt{3} + 1) + O(1/m)\sqrt{n}$$

Letting $m \to \infty$ we obtain the required result. []

APPENDIX TO CHAPTER III

This appendix collects proofs of the assertions in Chapter III. In contrast with Appendix II, proofs are informal. With additional work the results can be made rigorous - for example, Lemma 2.2 can be used as a model for Lemma 4.1 and Lemma 2.3 as a model for Theorem 4. However, it is believed that this would blur their simplicity.

3.1 The Static, Single-Bus Problem

Lemma 4.1

(i)
$$\lim_{n\to\infty} \frac{w_n(T_m^o)}{\sqrt{n}} = \frac{2}{15}\,[8\sqrt{2} - 7]\, b\sqrt{a} + O(1/m) \quad \text{a.e.}$$
$$= .5752 b\sqrt{a} + O(1/m) \quad \text{a.e.}$$

(ii)
$$\lim_{n\to\infty} \frac{f_n(T_m^o)}{\sqrt{n}} = \frac{1}{105}\,[118\sqrt{2} - 47]\, b\sqrt{a} + O(1/m) \quad \text{a.e.}$$
$$= 1.1417 b\sqrt{a} + O(1/m) \quad \text{a.e.}$$

Proof

Label the m regions $r_1, r_2, \ldots r_m$.

By Lemma 1.0, we can assume that, for any $i, j = 1, \ldots m$, there are n/m^2 demands of the form (r_i, r_j). Let

t_i be the time spent in r_i on the first passage

t_i' be the time spent in r_i on the second passage

$T = \sum_{i=1}^{m} t_i$ = time-to-completion of first passage.

For n large enough we have, with probability 1,

$$t_i = \frac{b\sqrt{n}}{m^{3/2}}\sqrt{m+i-1} \tag{A3.1}$$

since n/m origins and $(i-1)n/m^2$ destinations are collected. We can assume that the expected time at which any of these points is visited is $(t_1+t_2+\ldots+t_{i-1}+t_i/2)$.

Similarly,

$$t_i' = \frac{b\sqrt{n}}{m^{3/2}}\sqrt{m-i+1} \tag{A3.2}$$

since $(m-i+1)n/m^2$ destinations are visited in each region r_i. The expected time at which these points are visited is

$$T + t_1' + t_2' + \ldots + t_i'/2.$$

First, let us consider the waiting times. The sum of the waiting times in region i is

$$\frac{n}{m}\left[\sum_{j=1}^{i-1} t_j + t_i/2\right].$$

The total average waiting time is then

$$w_n(T_m^o) = \frac{1}{n}\frac{n}{m}\sum_{i=1}^{m}\left[\sum_{j=1}^{i-1} t_j + t_i/2\right]$$

$$= \frac{b\sqrt{n}}{m^{7/2}}\left[\sum_{i=1}^{m}(m-i+1/2)\sqrt{m+i-1}\right]$$

by expanding and using (A3.1).

Using a technique similar to that of Lemma 2.1, it can be shown that this series converges and

$$\frac{1}{m^{3/2}}\sum_{i=1}^{m}(m-i+1/2)\sqrt{m+i-1} = \frac{2}{15}[8\sqrt{2} - 7] + O(1/m)$$

$$= .5752 + O(1/m)$$

This proves Equation (i). Equation (ii) is somewhat

more cumbersome.

For destinations which are visited on the first passage, the sum of delivery times is

$$G_1 = \sum_{i=1}^{m} (i-1)\frac{n}{m^2}(t_1 + \dots + t_{i-1} + t_i/2)$$

$$= \frac{bn^{3/2}}{m^{7/2}} \frac{1}{2} \sum_{i=1}^{m} \sqrt{m+i-1}\,(m^2 - m - i^2 + 2i - 1)$$

Again it can be shown that

$$\frac{G_1}{n^{3/2}} = b\left[\frac{2}{5}(2^{5/2}-1) - \frac{1}{7}(2^{7/2}-1)\right] + O(1/m)$$

The sum of delivery times for passengers delivered on the second passage is

$$G_2 = \sum_{i=1}^{m} [T + t_1' + \dots + t_{i-1}' + t_i'/2]\,\frac{(m-i+1)n}{m^2}$$

$$= \frac{bn^{3/2}}{m^{7/2}}\left[\sum_{i=1}^{m} \sqrt{m+i-1}\,\frac{m}{2}(m+1) + \sum_{i=1}^{m} (m-i+1)^{5/2}\right]$$

Once again it can be shown that

$$\frac{G_2}{n^{3/2}} = b\left[\frac{1}{3}(2^{3/2}-1) + \frac{1}{7}\right] + O(1/m)$$

Finally, the average flow time is $(G_1+G_2)/n$, i.e.

$$\frac{f(T_m^o)}{\sqrt{n}} = b\left[\frac{2}{5}(2^{5/2}-1) - \frac{1}{7}(2^{7/2}-1) + \frac{1}{3}(2^{3/2}-1) + \frac{1}{7}\right] + O(1/m)$$

$$= \frac{b}{105}[118\sqrt{2} - 47] + O(1/m)$$

$$= 1.1417b + O(1/m).$$ []

The proof of Lemma 4.2 follows easily from Lemma 4.1 and the discussion in the text of Section 3.1.

<u>Theorem 4</u>

Let $f_n^* = \inf\{f_n(T),\ T \in S\}$.

Given any $\varepsilon > 0$, there exists an $M(\varepsilon)$ such that for $m \geq M$ and for n large enough, T_m^o is ε-optimal for problem (3.1) a.e.

Hence,

$$\lim_{n\to\infty} \frac{f_n^*}{\sqrt{n}} = 1.1417 b\sqrt{a} \qquad \text{a.e.}$$

<u>Proof</u>

Throughout this proof we assume that n is very large, so results obtained are asymptotic in n. For simplicity, we drop the parameterization on n.

For any problem instance, let the tour T_m^* minimize $\{f(T),\ T \in S_m\}$. Thus, for all m,

$$f(T_m^*) \leq f(T_m^o).$$

Let $\varepsilon > 0$ be given, and suppose that the hypothesis is not true. Then, for a subset W' of W with non-null measure, there exists M_1 such that for all $m \geq M_1$,

$$f(T_m^o) - f(T_m^*) > 3\varepsilon/2 \tag{A3.3}$$

By Lemma 4.3 there exists M_2 such that for all $m \geq M_2$,

$$w(T_m^o) + f(T_m^o) - w^+ + f^+ \leq \varepsilon/2 \quad \text{a.e.} \tag{A3.4}$$

So, for $m > \max\{M_1, M_2\}$,

$$\begin{aligned} f(T_m^*) + w(T_m^*) &\geq f^+ + w^+ && \text{a.e.} \\ &\geq w(T_m^o) + f(T_m^o) - \varepsilon/2 && \text{a.e. by (A3.4)} \\ &\geq w(T_m^o) + f(T_m^*) + \varepsilon && \text{a.e. by (A3.3)} \end{aligned}$$

i.e. $$w(T_m^*) - w(T_m^o) \geq \varepsilon \quad \text{a.e. in } W' \tag{A3.5}$$

For any partition of R into m equal subregions we can define a (nonfeasible) tour S_m^o that is similar to

T_m^o:- visit the regions successively, as for tour T_m^o, but in addition visit the destinations of the form (r_i, r_i) in each region r_i on the first passage through the m regions. Clearly,

$$w(S_m^o) > w(T_m^o).$$

But, by choosing m large enough - i.e. $m>M_3$, say - we have[1]

$$w(S_m^o) - w(T_m^o) < \varepsilon/2 \qquad (A3.6)$$

Let $M = \max\{M_1, M_2, M_3\}$. For any $m \geq M$ consider the tour T_m^* for any problem instance. We construct another partition of R. (This is again an extension of the method in Theorem 2; see that proof for notation.)

Divide the tour T_m^* *into two segments:* S' is the segment from the start of the tour to the delivery of the last origin; S'' is the remaining segment containing only destinations. Now divide S' into m subsegments: S_1' is the first segment from the first origin to the (n/m)th origin; S_i' is the (i)th segment from the $((i-1)n/m)$th origin to the (in/m)th origin. There exists a partition of R into m subregions $\{r_1, \ldots r_m\}$ such that

$$r_i^o \cap S_i' \neq \emptyset \qquad (r_i^o \text{ is the interior of } r_i)$$

$$r_i^o \cap S_j' = \emptyset \qquad \text{for } i \neq j$$

(Note that S' might cross itself, in which case different segments have points in common.)

---

(1) Observe that this is precisely the technique used in Theorem 2, Equation (A2.9a). The relationship (A3.6) can be justified in the same way; the formulas are particularly cumbersome, and would obscure the simplicity of the argument.

The space of all such partitions is nonempty. Let us draw a partition uniformly from this space, and consider the tours T_m^o and S_m^o respectively on <u>this</u> partition of R. We can assume that the origins and destinations are uniformly distributed within each subregion.

By construction,

$$w(T_m^*) \leq w(S_m^o).$$

So, by (A3.5) and (A3.6),

$$\begin{aligned} w(T_m^o) + \varepsilon/2 &> w(S_m^o) \\ &\geq w(T_m^*) \\ &> w(T_m^o) + \varepsilon \end{aligned}$$

which holds a.e. in W'. This implies that 1/2 > 1, a contradiction.

Thus, the first part of the hypothesis is true. By letting $m \to \infty$ we obtain, from Lemma 4.2,

$$\lim_{n\to\infty} \frac{f^*}{\sqrt{n}} = 1.1417b\sqrt{a} \quad \text{a.e.}$$

[]

<u>3.2 The Dynamic Single-Bus Problem</u>

<u>Lemma 5.1</u>

If the system is in steady-state, then with probability 1, the tour T_m^o given by Algorithm 5 minimizes, among tours in S_m, the average flow-time of all passengers.

<u>Proof</u>

Label the demand pairs as (o_i, d_i), $i=1,2,\ldots n,\ldots$ Let $t_1(i)$ be the time of arrival of (o_i, d_i); let $t(o_i)$ be

the time at which o_i is visited; and let $t(d_i)$ be the time at which d_i is visited. The average flow-time of the demands is then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} [t(d_i) - t_1(i)] \tag{A3.7}$$

Now regard the system as a queuing system (see e.g. [11]). It is convenient to consider two arrival streams - the origins $\{o_i\}$ and the destinations $\{d_i\}$. The arrival epoch of o_i is $t_1(i)$ and its departure epoch is $t(o_i)$. The arrival epoch of d_i is $t(o_i)$ and its departure epoch is $t(d_i)$. Thus, at any time t the 'units' to be 'served' which are present in the system are the points which can feasibly be visited at t.

The average flow-time of all <u>units</u> is

$$F = \lim_{n\to\infty} \frac{1}{2n} \{ \sum_{i=1}^{n} [t(o_i) - t_1(i)] + \sum_{i=1}^{n} [t(d_i) - t(o_i)]\}$$

$$= \lim_{n\to\infty} \frac{1}{2n} \sum_{i=1}^{n} [t(d_i) - t_1(i)] \tag{A3.8}$$

which is half the criterion we wish to minimize in (A3.7). So, minimizing (A3.8) will also minimize (A3.7).

We seek a stationary decision rule, X:- whenever a bus exits from a region, this rule will determine which region is to be visited next and which points are to be visited there. Any stationary such rule will induce a stationary arrival pattern on the destinations.

Given any rule X, let

N_X = average number of units in the system

T_X = average time between arrivals

F_X = average flow-time of units.

If the system is in steady-state, a well-known result of Little [13] gives that

$$F_X = T_X . N_X \quad \text{a.e.} \tag{A3.9}$$

Note that since X is stationary, the average time between arrivals of units is constant, i.e.

$$T_X = 1/2q \tag{A3.10}$$

(q is the constant arrival-rate of demands).

Further, if n units are visited in a region, then the average time spent on each unit is h_n/n, where h_n is the length of the path visiting these n points. For n large enough, h_n/n will be a decreasing function of n with probability 1. Now, to minimize N_X, the average number of units in the system, visit those which will take the shortest time - i.e. minimize h_n/n. Thus, choose the region with most points and, using an optimal travelling salesman tour, visit all points there. By (A3.9) and (A3.10) this decision rule will also minimize F_X (a.e.).

With (A3.8) the result is proven. []

__Theorem 5__

Let $F_{q,m}$ be the optimal average flow-time for tours in S_m when the arrival rate of passengers is q. Then,

$$\lim_{q\to\infty} \frac{F_{q,m}}{q} = 2b^2a(1+1/m) \quad \text{a.e.}$$

__Proof__

By Lemma 5.1 the tour T_m^o minimizes the average

flow-time in S_m, so we calculate the average flow-time of this tour.

Let θ be the time spent in each of the m equal areas(2). Then, by Equation (3.3),

$$\lim_{q\to\infty} \frac{\theta}{q} = 2b^2 \frac{a}{m} \quad \text{a.e.} \tag{A3.11}$$

Consider a time interval $[0,\theta]$ during which a particular region is visited. During this time $q\theta$ origins and $q\theta$ destinations are visited; we can take the average time-of-visitation as $t \approx \theta/2$. During each of the previous m intervals, of θ units each, $q\theta/m$ origins and $q\theta/m$ destinations arrived at this subregion - here by 'arrived' we refer to the arrival of units or points at the queuing system as described in Lemma 5.1. We can take the average arrival times of these $2q\theta/m$ points as $t_i = -(i-1/2)\theta$, for $i=1,2,...m$. Hence, the average flow-time of all points is

$$\frac{1}{2q\theta} \frac{2q\theta}{m} [\theta + 2\theta + ... + m\theta] = \frac{(m+1)\theta}{2}$$

By (A3.8) of Lemma 5.1, the average flow-time of passengers is twice this, i.e.

$$F_{q,m} = (m+1)\theta$$

and

$$\lim_{q\to\infty} \frac{F_{q,m}}{q} = 2b^2 a(1+1/m) \quad \text{a.e. by (A3.11).}$$

[]

---

(2) In the text of Section 3.2 we called this quantity θ_q.

3.3 Dynamic Multiple-Bus Problems

Theorem 6

The tour $T^k_{h;m}$ described by Algorithm 6 is optimal (asymptotically a.e.) for the k-bus minimum flow-time problem when tours are in S_{khm} and at most h transfers are allowed per period.

Let $F^k_{q;h,m}$ be the average flow-time resulting from this tour. Then,

$$\lim_{q\to\infty} \frac{F^k_{q;h,m}}{q} = \frac{b^2 a}{k^2}\left[2 + \frac{1}{mh}(1+1/k) + \frac{1}{h}(1-1/k)\right] \quad \text{a.e.} \qquad \text{(A3.12)}$$

Proof

We shall not prove the optimality of the tour $T^k_{h;m}$. A proof could proceed along the lines of Lemma 5.1, and the same principle would still be valid: each time that a bus exits from a region, visit next the region with most feasible points, and visit all feasible points there. For q large this will yield the tour of Algorithm 6.2. Further, with probability 1, each bus will spend the same time in each region, and there will be no waiting at the transfer point.

Below, we derive (A3.12) in detail.

Each bus visits m regions between any two transfers. Let θ be the time spent by a bus in any subregion. The period is then $hm\theta$. During this time there are $qhm\theta$ new demands, and so each bus must serve $2qhm\theta/k$ points. In each region there are $2qhm\theta/khm = 2q\theta/k$ points visited.

Hence,

$$\theta = b\sqrt{\frac{2q\theta}{k}}\sqrt{\frac{a}{khm}}$$

i.e.
$$\theta = \frac{2b^2qa}{k^2hm} \tag{A3.13}$$

Each time a region is visited $q\theta/k$ origins and $q\theta/k$ destinations are visited.

Consider a bus B visiting a region that is the j-th after a transfer point. Of the $q\theta/k$ destinations, $q\theta/k^2$ were collected (i.e. their corresponding origins were visited) by bus B itself, and $(k-1)q\theta/k^2$ were collected by the $(k-1)$ other buses. The $q\theta/k^2$ points were visited uniformly during the last hm time intervals, of length θ each. So the waiting-time for these points is

$$\frac{q\theta}{k^2hm}(\theta + 2\theta + \ldots + hm\theta)$$
$$= \frac{q\theta}{2k^2}(hm+1)\theta \tag{A3.14}$$

The $(k-1)q\theta/k^2$ points were collected uniformly during the hm time intervals before the last transfer, $(j-1)$ intervals previously. Thus, the total waiting-time for these points is

$$\frac{(k-1)q\theta}{k^2hm}[j\theta + (j+1)\theta + (j+2)\theta + \ldots + (hm+j-1)\theta]$$
$$= \frac{(k-1)q\theta}{k^2}(hm+2j-1)\theta \tag{A3.15}$$

The $q\theta/k$ origins all arrived during the last hm time intervals. The total waiting-time for these points is

$$\frac{q\theta}{khm}(\theta + 2\theta + \ldots + hm\theta)$$
$$= \frac{q\theta}{2k}(hm+1)\theta \tag{A3.16}$$

The total waiting-time for all points visited during this j-th region of bus B is the sum of (A3.14), (A3.15) and (A3.16). This is

$$\frac{q\theta}{2k^2}.\theta\ [(hm+1)(k+1) + (hm+2j-1)(k-1)]$$

and $2q\theta/k$ points were visited.

Summing now for $j=1,2,\ldots m$ and dividing by $2q\theta m/k$ we obtain the average waiting-time of all points, viz.

$$\frac{\theta}{4km}\ [m(hm+1)(k+1) + \sum_{j=1}^{m} (hm+2j-1)(k-1)]$$

$$= \frac{\theta m}{4km}\ [khm+k+hm+1 + khm-mh+mk-m]$$

$$= \frac{\theta m}{4k}\left[2kh + \frac{(k+1)}{m} + (k-1)\right]$$

For the required average flow-time of passengers we must multiply this by 2. Also, using (A3.13) we get

$$F^{k}_{q;h,m} = \frac{1}{2k}\,\frac{2b^2qa}{k^2h}\left[2kh + \frac{(k+1)}{m} + (k-1)\right]$$

$$= \frac{b^2qa}{k^2}\left[2 + \frac{1}{mh}(1+1/k) + \frac{1}{h}(1-1/k)\right]$$

This gives the required result. []

REFERENCES

1. Aho, A.V., J.E. Hopcroft and J.D. Ullman, 'The Design and Analysis of Computer Algorithms', Addison-Wesley, Reading, 1974.

2. Beardwood, J., J.H. Halton and J.M. Hammersley, The Shortest Path through Many Points, Proc. Cambridge Philosophical Society, 55, pp 299-327, 1959.

3. Christofides, N., Worst-Case Analysis of a New Heuristic for the Travelling Salesman Problem, Management Sciences Research Report NO. 388, Carnegie-Mellon University, Pittsburgh, Pennsylvania.

4. Christofides, N., and S. Eilon, Expected Distances in Distribution Problems, Operational Research Quarterly, 20, pp 437-443, 1969.

5. Coffman, E.G., (ed.), 'Computer and Job-Shop Scheduling Theory', Wiley-Interscience, New York, 1976.

6. Conway, R.W., W.L. Maxwell and L.W. Miller, 'Theory of Scheduling', Addison-Wesley, Reading, 1967.

7. Dreyfus, S.E., 'Dynamic Programming and the Calculus of Variations', Academic Press, New York, 1965.

8. Held, M., and R.M. Karp, A Dynamic Programming Approach to Sequencing Problems, J. Soc. Indust. Appl. Math., 10, pp 196-210, 1962.

9. Karp, R.M., The Probabilistic Analysis of Some Combinatorial Algorithms, Memo. No. ERL-M58, University of California, Berkeley, 1976.

10. Kates, J., 'Para Transit Dispatch Systems', J. Kates Associates Inc., Toronto, Ontario, 1975.

11. Kleinrock, L., 'Queuing Systems: Vol. 1, Theory', Wiley, New York, 1975.

12. Lin, S., and B.W. Kernighan, An Effective Heuristic Algorithm for the Travelling Salesman Problem, Operations Research, 21, pp 498-516, 1973.

13. Little, J.D.C., A Proof of the Queuing Formula: L= W, Operations Research, 9, pp 383-387, 1961.

14. Loeve, M., 'Probability Theory', Van Nostrand, Princeton, 1963.

15. Newell, G.F., Control of Pairing of Vehicles on a Public Transportation Route: Two Vehicles, One Control Point, Transportation Science, 8, pp 248-264, 1974.

16. Papadimitriou, C.H., The Euclidean Travelling Salesman Problem is NP-Complete, J. Theoretical Computer Science, to appear.

17. Potter, B., Ann Arbor's Dispatching System, presented at the Sixth International Conference on Demand-Responsive Transportation Systems and Other Paratransit Services, March 1976.

18. Rebibo, K.K., A Computer Controlled Dial-a-Ride System, in AFCET (ed.), Traffic Control and Transportation Systems, North Holland Publishing Co., Amsterdam, 1974.

19. Rosenbloom, S., (ed.), 'Paratransit', Special Report 164, Transportation Research Board, Washington D.C., 1976.

20. Royden, H.L., 'Real Analysis', Macmillan, New York, 1963.

21. Wilson, N.H.M., et. al., Scheduling Algorithms for a Dial-a-Ride System, Massachusetts Institute of Technology, Urban Systems Laboratory, Technical Report USL TR-70-13, March 1971.

22. Wilson, N.H.M., et. al., Advanced Dial-a-Ride Algorithms Research Project: Final Report, Massachusetts Institute of Technology, Dept. of Materials Science and Engineering, Technical Report R76-20, March 1976.

JSEP REPORTS DISTRIBUTION LIST
N00014-75-C-0648

Defense [illegible] Center
Attn: DDC-TCA (Mrs. V. Caponio)
Cameron Station
Alexandria, Virginia 22314

Asst. Director
Electronics and Computer Sciences
Office of Director of Defense Research
and Engineering
The Pentagon
Washington, DC 20315

Office of the Director of Defense
Research and Engineering
Information Office Library Branch
The Pentagon
Washington, DC 20301

ODDR and E Advisory Group on
Electron Devices
201 Varick Street
New York, New York 10014

Chief, R and D Division (340)
Defense Communications Agency
Washington, DC 20305

Director, National Security Agency
Fort George G. Meade, Maryland 20755
Attn: Dr. T. J. Beahn

Institute for Defense Analysis
Science and Technology Division
400 Army Navy Drive
Arlington, Virginia 22202

Dr. Buchley
Defense Advanced Research Projects Agency
Attn: Technical Library
1400 Wilson Boulevard
Arlington, Virginia 22209

Dr. R. Reynolds
Defense Advanced Research Projects Agency
Attn: Technical Library
1400 Wilson Boulevard
Arlington, Virginia 22209

AF/RDPS
The Pentagon
Washington, DC 20330

AFSC (LJ/Mr. Irving R. Mirman)
Andrews AFB
Washington, DC 20334

Directorate of Electronics and Weapons
HQ AFSC/DLC
Andrews AFB, Maryland 20334

Directorate of Science
HQ AFSC/DLS
Andrews AFB
Washington, DC 20334

LTC J. W. Gregory
AF Member, TAC
Air Force Office of Scientific Research
Bolling AFB
Washington, DC 20332

Mr. Carl Sletten
AFCRL/LZ
Hanscom AFB, MA 01731

Dr. Richard Picard
AFCRL/OPL
Hanscom AFB, MA 01731

Mr. Robert Barrett
AFCRL/LQ
Hanscom AFB, MA 01731

Dr. John N. Howard
AFCRL (CA)
Hanscom AFB, MA 01731

Dr. Richard B. Mack
AFCRL/LZR
Hanscom AFB, MA 01731

Documents Library (TILD)
Rome Air Development Center
Griffiss AFB, New York 13441

Mr. H. E. Webb, Jr. (ISCP)
Rome Air Development Center
Griffiss AFB, New York 13441

Mr. Murray Kesselman (ISCA)
Rome Air Development Center
Griffiss AFB, New York 13441

Mr. W. Edwards
AFAL/TE
Wright-Patterson AFB, Ohio 45433

Mr. R. D. Larson
AFAL/DHR
Wright-Patterson AFB, Ohio 45433

Chief Scientist
AFAL/CA
Wright-Patterson AFB, Ohio 45433

HQ ESD (DRI/Stop 22)
Hanscom AFB, MA 01731

Professor R. E. Fontana
Head Dept. of Electrical Eng
AFIT/ENE
Wright-Patterson AFB, Ohio 45433

Mr. John Mottsmith (MCIT)
HQ ESD (AFSC)
Hanscom AFB, MA 01731

LTC Richard J. Gowen
Professor
Dept. of Electrical Engineering
USAF Academy, Colorado 80840

AUL/LSE-9663
Maxwell AFB, Alabama 36112

AFETR Technical Library
P.O. Box 4608, MU 5650
Patrick AFB, Florida 32542

ADTC (DLOSL)
Eglin AFB, Florida 32542

HQ AMD (RDR/COL Godden)
Brooks AFB, Texas 78235

USAF European Office of Aerospace
Research
Technical Information Office
Box 14, FPO, New York 09510

Dr. Carl E. Baum
AFWL (ES)
Kirtland AFB, New Mexico 87117

USAFSAM/RAL
Brooks AFB, Texas

HQDA (DAMA-ARZ-A)
Washington, DC 20310

Commander
US Army Security Agency
Attn: IARD-T
Arlington Hall Station
Arlington, VA 22212

Commander
US Army Material Development
and Readiness Command
Attn: Technical Library, Rm 7B 35
5001 Eisenhower Avenue
Alexandria, VA 22333

Commander
US Army Ballistics Research
Laboratory
Attn: DRXRD-BAD
Aberdeen Proving Ground
Aberdeen, MD 21005

Commander
Picatinny Arsenal
Attn: SMUPA-TS-T-S
Dover, NJ 07801

US Army Research Office
Attn: Dr. Hermann Robl
P.O. Box 12211
Research Triangle Park, NC 27709

US Army Research Office
Attn: Mr. Richard O. Ulsh
P.O. Box 12211
Research Triangle Park, NC 27709

US Army Research Office
Attn: Dr. Jimmie R. Suttle
P.O. Box 12211
Research Triangle Park, NC 27709

US Army Research Office
Attn: Dr. Horst Wittmann
P.O. Box 12211
Research Triangle Park, NC 27709

Commander
Frankford Arsenal
Attn: Mr. George C. White, Jr.
Deputy Director, Pitman-Dunn Lab.
Philadelphia, PA 19137

Commander
US Army Missile Command
Attn: Chief, Document Section
Redstone Arsenal, AL 35809

Commander
US Army Missile Command
Attn: DRSMI-RR
Redstone Arsenal, AL 35809

Commander
US Army Materials and Mechanics
Research Center
Attn: Chief, Materials Sciences
Division
Watertown, MA 02172

Commander
Harry Diamond Laboratories
Attn: Mr. John E. Rosenberg
2800 Powder Mill Road
Adelphi, MD 20783

Commandant
US Army Air Defense School
Attn: ATSAD-T-CSM
Fort Bliss, TX 79916

Commandant
US Army Command and General
Staff College
Attn: Acquisitions, Lib. Div.
Fort Leavenworth, Kansas 66027

Dr. Hans K. Ziegler
Army Member, TAC/JSEP
US Army Electronics Command
(DRSEL-TL-D)
Fort Monmouth, NJ 07703

Mr. J. E. Teti
Executive Secretary, TAC/JSEP
US Army Electronics Command
(DRSEL-TL-DT)
Fort Monmouth, NJ 07703

Director
Night Vision Laboratory, ECOM
Attn: DRSEL-NV-D
Fort Belvoir, VA 22060

Commander/Director
Atmospheric Sciences Laboratory
(ECOM)
Attn: DRSEL-BL-DD
White Sands Missile Range, NM 88002

Director
Electronic Warfare Laboratory
(ECOM)
Attn: DRSEL-WL-MY
White Sands Missile Range, NM 88002

Commander
US Army Armament Command
Attn: DRSAR-RD
Rock Island, IL 61201

Director, Div. of Neuropsychiatry
Walter Reed Army Institute of
Research
Washington, DC 20012

Commander
USASATCOM
Fort Monmouth, NJ 07703

Commander
US Army R and D Group
(Far East)
APO San Francisco, CA 96343

Commander
US Army Communications Command
Attn: Director, Advanced Concepts
Office
Fort Huachuca, AZ 85613

Project Manager
ARTADS
EAI Building
West Long Branch, NJ 07764

Commander
US Army White Sands Missile Range
Attn: STEWS-ID-R
White Sands Missile Range, NM 88002

Director, TRI-TAC
Attn: TT-AD (Mrs. Briller)
Fort Monmouth, NJ 07703

Commander
US Army Communication Command
Attn: CC-OPS-PD
Fort Huachuca, AZ 85613

COL Robert Noce
Senior Standardization Representative
US Army Standardization Group,
Canada
Canadian Force Headquarters
Ottawa, Ontario, Canada K1A OK2

Commander
US Army Electronics Command
Attn: DRSEL-RD-O (Dr. W. S. McAfee)
CT-L (Dr. R. Buser)
NL-O (Dr. H. S. Bennett)
NL-T (Mr. R. Kulinyi)
TL-B
VL-D
WL-D
TL-MM (Mr. N. Lipetz)
NL-H (Dr. F. Schwering)
TL-E (Dr. S. Kronenberg)
TL-E (Dr. J. Kohn)
TL-I (Dr. C. Thornton)
NL-B (Dr. S. Amoroso)
Fort Monmouth, NJ 07703

Project Manager
Ballistic Missile Defense Program Office
Attn: DACS-BMP (Mr. A. Gold)
1300 Wilson Blvd.
Washington, DC 22209

Dr. Sidney Ross
Technical Director
SARFA-TD
Frankford Arsenal
Philadelphia, PA 19137

Department of Naval Research
Electronic and Solid State
Sciences Program (Code 427)
800 N. Quincy
Arlington, Virginia 22217

Office of Naval Research
Code 200
Assistant Chief for Technology
800 N. Quincy
Arlington, Virginia 22217

Office of Naval Research
Information Sciences Program
(Code 437)
800 N. Quincy
Arlington, Virginia 22217

Naval Research Laboratory
4555 Overlook Avenue, S.W.
Washington, DC 20375
Attn: Codes 2627
4000
4105
5000
5200
5203
5210
5270
5300
5400
5460
5464
5500
5510
6400

Director
Office of Naval Research Branch Office
536 South Clark Street
Chicago, Illinois 60605

San Francisco Area Office
Office of Naval Research
760 Market Street, Room 447
San Francisco, CA 94102

Dr. A. Laufer
Chief Scientist
Office of Naval Research Branch Office
1030 East Green Street
Pasadena, CA 91101

Director
Office of Naval Research
Branch Office
715 Broadway, 5th Floor
New York, New York 10003

Mr. L. W. Sumney
Naval Electronics Systems Command
NC No. 1
2511 Jefferson Davis Highway
Arlington, VA 20360

Mr. B. Frotin
Naval Electronics Systems Command
NC No. 1
2511 Jefferson Davis Highway
Arlington, VA 20360

Mr. N. Butler
Naval Electronics Systems Command
NC No. 1
2511 Jefferson Davis Highway
Arlington, VA 20360

Dr. H. J. Mueller
Naval Air Systems Command
JP No. 1
1411 Jefferson Davis Highway
Arlington, VA 20360

Capt. R. B. Meeks
Naval Sea Systems Command
NC No. 3
2531 Jefferson Davis Highway
Arlington, VA 20362

Commander
Naval Surface Weapons Center
Attn: Technical Library
Silver Spring, Maryland 20910

Officer-in-Charge
Naval Surface Weapons Center
Dahlgren Laboratory
Dahlgren, Virginia 22448

Naval Air Development Center
Attn: Technical Library
Johnsville
Warminster, Penn. 18974

Commander
Naval Avionics Facility
Indianapolis, Indiana 46241
Attn: D/035 Technical Library

Naval Missile Center
Technical Library
Code 5632.2
Point Mugu, CA 93042

Naval Weapons Center
Attn: Technical Library
Code 533
China Lake, CA 93555

Naval Weapons Center
Attn: Code 6010
China Lake, CA 93555

Naval Weapons Center
Mathematics Division
China Lake, CA 93555

Naval Training Equipment Center
Technical Library
Orlando, Florida 32813

Naval Research Laboratory
Underwater Sound Reference Div.
Technical Library
P.O. Box 8337
Orlando, Florida 32806

Naval Underwater Sound Center
Technical Library
New London, Conn. 06320

U.S. Naval Oceanographic Office
Library - Code 1600
Washington, DC 20373

Commandant, Marine Corps
Scientific Advisor (Code AX)
Washington, DC 20380

Dr. Gernot M. R. Winkler
Director, Time Service
US Naval Observatory
Mass. Ave. at 34th St. N.W.
Washington, DC 20390

Naval Postgraduate School
Technical Library
Monterey, California 93940

Naval Electronics Laboratory Center
Technical Library
San Diego, California 92152

Naval Electronics Laboratory Center
Attn: Code 2600
San Diego, CA 92152

Naval Electronics Laboratory Center
San Diego, California 92152
Attn: Codes 2000
2200
2300
3000
3200
3300
4000
5000
5200
5300

Naval Undersea Center
Technical Library
San Diego, California 92152

Naval Ship Research and Development
Center
David W. Taylor
Code 522.1
Bethesda, Maryland 20084

Office of Chief of Naval Operations
NAICOM/MIS Planning Branch
NOP-916D, Pentagon
Washington, DC 20350

Mr. F. C. Schwenk, RD-T
National Aeronautics and
Space Administration
Washington, DC 20546

Los Alamos Scientific Laboratory
Attn: Reports Library
P.O. Box 1663
Los Alamos, New Mexico 87544

M. Zane Thornton
Deputy Director Institute for
Computer Sciences and Technology
National Bureau of Standards
Washington, DC 20234

Director, Office of Postal
Technology (R and D)
US Postal Service
11711 Parklawn Drive
Rockville, Maryland 20852

NASA Lewis Research Center
Attn: Library
21000 Brookpark Road
Cleveland, Ohio 44135

Library - R51
Bureau of Standards
Acquisition
Boulder, Colorado 80302

MIT Lincoln Laboratory
Attn: Library A-082
P.O. Box 73
Lexington, Mass. 02173

Dr. Jay Harris
Program Director
Devices and Waves Program NSF
1800 G Street
Washington, DC 20550

Dr. Howard W. Etzel
Deputy Director
Div. of Materials Research NSF
1800 G Street
Washington, DC 20550

Dr. Dean Mitchell
Program Director
Solid-State Physics
Div. of Materials Research
National Science Foundation
1800 G Street
Washington, DC 20550

Director
Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Director
Microwave Research Institute
Polytechnic Institute of New York
Long Island Graduate Center, Route 110
Farmingdale, New York 11735

Asst. Director
Microwave Research Institute
Polytechnic Institute of New York
333 Jay Street
Brooklyn, New York 11201

Director
Columbia Radiation Laboratory
Dept. of Physics
Columbia University
538 West 120th Street
New York, New York 10027

Director
Coordinated Science Laboratory
University of Illinois
Urbana, Illinois 61801

Director
Stanford Electronics Laboratory
Stanford University
Stanford, California 94305

Director
Microwave Laboratory
Stanford University
Stanford, California 94305

Director
Electronics Research Laboratory
University of California
Berkeley, California 94720

Director
Electronics Sciences Laboratory
University of Southern California
Los Angeles, California 90007

Director
Electronics Research Center
The University of Texas at Austin
Engineering-Science Bldg. 112
Austin, Texas 78712

Director of Laboratories
Div. of Engineering and Appl. Physics
Harvard University
Pierce Hall
Cambridge, Mass. 02138